# AN EXPLORATION OF NESTED RECURRENCES USING EXPERIMENTAL MATHEMATICS

BY NATHAN HAREL FOX

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Mathematics

Written under the direction of

Doron Zeilberger

And approved by

New Brunswick, New Jersey

May, 2017

## ABSTRACT OF THE DISSERTATION

# An Exploration of Nested Recurrences Using Experimental Mathematics


**by NATHAN HAREL FOX**

**Dissertation Director: Doron Zeilberger**



Broadly speaking, Experimental Mathematics is the philosophy that computers are a valuable tool that should be used extensively in mathematical research. Here, we apply this philosophy to the study of integer sequences arising from nested recurrence relations. The most widely studied nested recurrence is the Hofstadter $Q$-recurrence: $Q(n) = Q(n - Q(n - 1)) + Q(n - Q(n - 2))$. Hofstadter considered this recurrence with the initial condition $Q(1) = Q(2) = 1$, and the resulting sequence has much apparent structure. But, almost nothing has been rigorously proved about it. Others have modified the recurrence, the initial conditions, or both, to obtain related but more predictable sequences. We follow that vein and prove a number of unrelated theorems about sequences resulting from nested recurrences. Our first results relate to automatically finding (with proof) solutions to nested recurrences that are interleavings of linear-recurrent sequences. We then present a new nested-recurrent sequence whose terms increase monotonically with successive differences zero or one. Finally, we embark on an exploration of strange but predictable behaviors that result when recurrences are given various types of initial conditions.

## Acknowledgements

I would first like to thank my advisor, Doron Zeilberger. In coming to Rutgers, I knew that I liked using computers to attack mathematical problems, but I did not know that there was an entire industry devoted to such things. I quickly began attending the Experimental Mathematics Seminar and interacting with Dr. Zeilberger. His Spring 2013 Experimental Math course on games led to my first published results at Rutgers. He also introduced me to the work of Golomb [15] and Ruskey [30] on nested recurrences, which evolved into most of the results in this dissertation.

I also want to thank the other members of my thesis committee, Swastik Kopparty, János Komlós, and Neil Sloane, for their support throughout the years and their valuable feedback on this dissertation.

I have been lucky over the last five years to be surrounded by a friendly and strongly supportive community of graduate students. Thank you especially to the Rutgers Applied Game Theory Seminar group, and to Jake Baron, who started this group shortly before I arrived. I later accepted his transfer of responsibility for administering this group, and I am glad it has continued to thrive. I have also become a part of a legendary community in Highland Park.

I would specifically like to thank Matthew Russell, Nathaniel Shar, Anthony Zaleski, and Richard Voepel for reading drafts of some of my papers and providing valuable feedback. I also want to thank Nathaniel for having Douglas Hofstadter as an uncle. This led to Hofstadter speaking at the Experimental Math Seminar about his $Q$-recurrence [19], which led to my research interests in the field.

And, of course, my family has been nothing but supportive of me as I have pursued my doctorate.

Some portions of this dissertation have been published in various journals. Chapter 4

has been published as [12], and Chapter 5 has been published as [11]. In addition, various pieces of Chapter 2 and of Chapter 11 have appeared in introductory or concluding remarks in some of these publications, and other chapters are in various stages of the submission process. Preliminary versions of Chapter 3 and Chapter 6 plus Section 9.4 are available on ArXiv [10, 13]. For those documents that have gone through the peer review process, I would like to thank the various editors and anonymous reviewers.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

The techniques we use to arrive at the results in this dissertation are just as important as the results themselves. We employ methodologies that are firmly rooted in the philosophy of Experimental Mathematics, which emphasizes the importance of using computers in mathematical research [2]. Computers serve multiple roles in mathematical pursuits. First and most straightforwardly, computers can explore mathematical problems of interest and generate data about them. When trying to prove a result, it is quite reassuring if it has a large base of numerical evidence in its favor. Plus, the computer can be used to gather data in search *of* a conjecture, typically by exploring data in an attempt to determine which problems are worth studying further. Once noteworthy patterns are discovered, computers can frequently automate otherwise tedious inductive proofs or case analyses. Computers are also good at falsifying conjectures, as they can quickly carry out a search for a counterexample. Finally, computers and the Internet make invaluable databases like the Online Encyclopedia of Integer Sequences [31] possible.

The mathematical content of this dissertation consists of numerous results about nested recurrence relations. Recurrence relations occur frequently in many areas of mathematics. The simplest are homogeneous linear recurrence relations, which are of the form

$$a(n) = \sum_{i=1}^{k} \lambda_i a(n-1)$$

for some parameters $k \geq 1$ and $\lambda_1, \lambda_2, \ldots, \lambda_k$. Such recurrences are well understood, as are their nonhomogeneous counterparts. On the other hand, a general theory of nonlinear recurrences is less tractable, largely because of the diversity of possible forms and behaviors. Nonlinear recurrences frequently arise in the context of dynamical systems, where a sequence is generated by repeatedly iterating a given map. Such recurrences are typically highly sensitive to initial conditions. For example, the famously difficult $3x+1$ problem can be phrased in terms of an integer-valued dynamical system, and hence, as a question about nonlinear recurrences [6, 27]. As generating terms of a sequence defined by a nonlinear recurrence is typically straightforward, problems involving them naturally lend themselves to exploration via Experimental Mathematics.

Within the class of nonlinear recurrences, a particularly finicky subfamily are the *nested recurrences*, such as $a(n) = a(a(n-1))$. In this typical example, the previous terms that $a(n)$ depends on themselves depend on other sequence terms. The most widely studied nested recurrence is the Hofstadter $Q$-recurrence [18], which is defined by the nested recurrence

$$Q(n) = Q(n - Q(n-1)) + Q(n - Q(n-2)).$$

Most of our study focuses on this recurrence and generalizations of it.

The most common questions about sequences arising from nested recurrences are the following:

1. Does the sequence actually exist?

2. Can we efficiently compute the $n^{th}$ term of the sequence?

3. Is there a closed form for the terms of the sequence?

4. Does the sequence have any notable global properties?

5. Does the sequence have a combinatorial interpretation?

Throughout this dissertation, we are primarily concerned with items 2, 3, and 4, especially item 4. We consider two types of global properties. The first property, known as *slowness*, has been studied primarily by Tanny and various coauthors [3, 9, 20, 28, 32]. Slow sequences frequently arise from slight modifications of the Hofstadter $Q$-recurrence with simple initial conditions. The second global property is eventual satisfaction of a linear recurrence, or, equivalently, having a rational ordinary generating function [15, 30]. Prior to this work, all known examples of such sequences satisfied the

Hofstadter $Q$-recurrence itself with somewhat complicated initial conditions. We shall see that two sequences that appear completely unrelated oftentimes eventually satisfy the same nested recurrence. Such observations lead to the following mantra on nested recurrences: "If it seems like it might be possible, it probably is possible."

This thesis can broadly be thought of as consisting of two major "chunks." The first half consists of items closely related to results in the literature. Chapter 2 contains some definitions and more in-depth background about nested recurrences and the sequences they generate. Then, Chapter 3 describes a way to use symbolic computation to discover a multitude of solutions that satisfy linear recurrences. Explorations using this process lead to the discoveries of Chapters 4 and 5. Next, Chapter 6 contains a description of a previously unknown slow solution to a Hofstadter-like recurrence.

The second half of this dissertation presents a novel way of studying nested recurrences. The results of Chapters 3, 4, and 5 involve using the form of the solution to discover an initial condition. In Chapters 8, 9 and 10, we flip this around and use the form of the initial condition to discover the behaviors of families of sequences. Some of the results in these chapters are truly bizarre. For instance, one of our results (Theorem 9.3) may be best thought of as a result about sequences over 5-adic integers. And, another theorem (Theorem 9.8) characterizes all sequences of a certain type, except for 6081 special cases. One of these special cases (p. 134) is a finite sequence with $84975 \cdot 2^{560362} + 31$ terms.

Chapter 7 serves as a short transition between the two chunks. It is a fairly self-contained discussion of an unusual sort of solution to the Hofstadter $Q$-recurrence. At the end of the dissertation, Chapter 11 contains a summary of open problems and future research directions.

Much of the work of this dissertation uses programs written in Maple, and there are many computer generated proofs and data that are too lengthy and tedious to appear in this document. These supplemental materials can be found at `http://github.com/nhf216/thesis`, and they are all listed in Appendix G.

# Chapter 2

# Background on Nested Recurrences

A *recurrence relation* (often shortened to *recurrence*) is any definition of a sequence $(a(n))_{n\geq 1}$ by a rule defining $a(n)$ in terms of $n$ and previous values in the sequence. A recurrence relation alone is (typically) not enough to specify an actual sequence, as the first value (or, more commonly, the first few values) has no previous values to be defined in terms of. To actually obtain a sequence from a recurrence relation, some of the first values must be specified separately. Such additional definitions are together known as an *initial condition*. Throughout this dissertation, we denote the initial condition $a(1) = a_1$, $a(2) = a_2$, $\ldots$, $a(k) = a_k$ by $\langle a_1, a_2, \ldots, a_k\rangle$. Also, for convenience and consistency, we index all of our sequences from 1. Sequences can theoretically begin at any index, and our definitions and results can be adjusted accordingly.

When we use the term *solution* to describe a (finite or infinite) integer sequence, we mean that the terms of the sequence eventually satisfy a particular recurrence. All solutions that we consider in this dissertation are *explicit* solutions; that is, given the first $n-1$ terms, the recurrence relation can be used to generate the $n^{th}$ term in the sequence. (This is in contrast to *implicit* solutions, which may satisfy a recurrence without being generated by it.)

The simplest recurrence relations are homogeneous linear recurrences, which are of the form

$$a(n) = \sum_{i=1}^{k} \lambda_i a(n-1)$$

for some parameter $k \geq 1$ (called the *order*) and parameters $\lambda_1, \lambda_2, \ldots, \lambda_k$. We call a sequence *linear-recurrent* if its terms eventually satisfy some linear recurrence. The most well-known linear recurrence is the Fibonacci recurrence $F(n) = F(n-1) + F(n-2)$. With the initial condition $\langle 1, 1\rangle$, this recurrence generates the Fibonacci sequence [31, A000045], and with the initial condition $\langle 1, 3 \rangle$ it generates the Lucas sequence [31, A000032]. The Fibonacci numbers famously have the ordinary generating function

$$\sum_{n=1}^{\infty} F(n)x^n = \frac{x}{1 - x - x^2}$$

and the closed form

$$F_n = \frac{1}{\sqrt{5}} \left( \frac{1+\sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left( \frac{1-\sqrt{5}}{2} \right)^n .$$

In general, a sequence is linear-recurrent if and only if it has a rational generating function. Also, adding a polynomial or exponential nonhomogeneous part to a linear recurrence does not increase the power of such recurrences; if a sequence eventually satisfies a nonhomogeneous linear recurrence, it also satisfies a homogeneous linear recurrence of higher order.

Many real-world processes are conveniently modeled by recurrence relations. Unfortunately, most recurrences appearing in applications are *nonlinear*. In this thesis, we study a particular type of nonlinear recurrence: nested recurrences. A nested recurrence is a type of recurrence relation where the previous terms that the $n^{th}$ term depends on are not fixed in advance. Instead, they are themselves dependent on previous sequence terms. We call a sequence generated by a nested recurrence *nested-recurrent.* A general theme of this dissertation is that nested recurrence relations are highly sensitive to their initial conditions.

The most notable and well studied nested-recurrent sequence is the Hofstadter $Q$-sequence. It was first introduced by Douglas Hofstadter in the 1960s [18], and it is defined by the recurrence

$$Q(n) = Q(n - Q(n-1)) + Q(n - Q(n-2))$$

and the initial condition $\langle 1, 1 \rangle$. This definition superficially resembles that of the Fibonacci sequence, the disparity being that the previous two terms in the $Q$-sequence tell how far to go back to find two *other* terms to add together. For this reason, the $Q$-sequence is known as a meta-Fibonacci sequence. The behavior of the $Q$-sequence is nothing like the Fibonacci sequence, however. The first fifteen terms of the $Q$-sequence

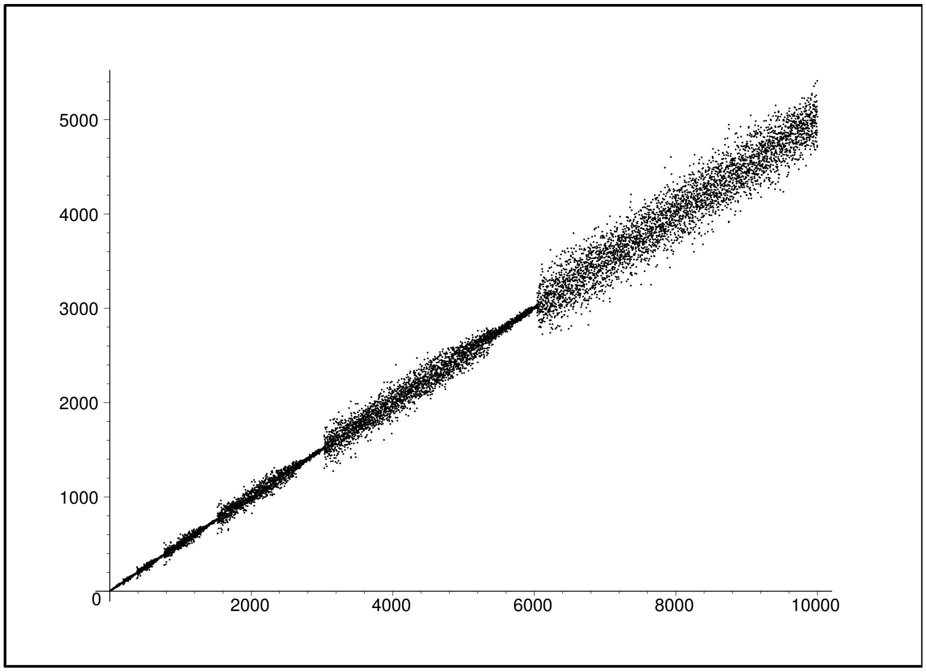


Figure 2.1: The first 10000 terms of the Hofstadter $Q$-sequence (A005185 in OEIS)

increase monotonically, but thereafter the sequence rapidly devolves into chaos. On the global level, $Q(n)$ seems to oscillate around $\frac{n}{2}$, and a plot of the sequence appears to be built of sausage-like structures, each one twice as large as the previous (see Figure 2.1). But, all that is known rigorously is that if

$$\lim_{n\to\infty} \frac{Q(n)}{n}$$

exists, it must equal one half [15]. Furthermore, it is unknown whether $Q(n)$ even *exists* for all $n$. If $Q(n-1) \geq n$ for some $n$, then $Q(n)$ would be defined in terms of $Q(k)$ for some $k \leq 0$. But, $Q$ is not defined on nonpositive inputs, so $Q(n)$ would fail to exist. All subsequent terms would also fail to exist, so the sequence would be finite in this scenario. If a sequence is finite in this manner, the sequence is said to *die*. We do not know for sure whether the Hofstadter $Q$-sequence dies [16]. We do know that it exists for at least $10^{10}$ terms [31, A005185], and empirical evidence suggests that it lives forever [29]. In general, a sequence that dies may live for a long time before it dies. In fact, the death question is undecidable given a generic nested recurrence and initial condition [5].

There are actually two ways in which a sequence can die. The first and most common way is what we saw before: if the $n^{th}$ term depends on some term from before the initial condition. The other possibility is that the $n^{th}$ term depends on itself or some future term. (Dependence on a future term may not be an issue when considering implicit solutions.) For us, the classical notion of death is overly restrictive, as it requires any solution to any nested recurrence with a term of the form $U(n - U(n-1))$ to grow no faster than $n$. So, going forward, we have two notions of death.

**Definition 2.1.** *A sequence generated by a nested recurrence $U$* weakly dies *if, in order to compute $U(n)$, the value of $U(m)$ is needed, where $m \leq 0$ or $m \geq n$.*

The notion of weak death is identical to the notion of death we have already discussed.

**Definition 2.2.** *A sequence generated by a nested recurrence $U$* strongly dies *if, in order to compute $U(n)$, the value of $U(m)$ is needed where $m \geq n$. If a value $U(m)$ with $m \leq 0$ is needed, use $0$ for that value.*

In other words, we "cheat" and define $U(n) = 0$ for all $n \leq 0$. This convention, which has been used by other authors [30], allows us to observe a wide variety of solutions that grow faster than $n$.

A nested-recurrent sequence can have any of the following eventual behaviors:

**Definition 2.3.**

- *If a sequence strongly dies and weakly dies at the same index, then it* vehemently dies.
- *If a sequence strongly dies at some index and weakly dies at the same or some earlier index, then it* strongly dies. *(In this way, saying that a sequence strongly dies does not preclude it vehemently dying.)*
- *If a sequence weakly dies but does not strongly die, then it* persists.
- *If a sequence does not weakly die, then it* lives.

We know that the Hofstadter $Q$-sequence does not vehemently die, as there is no way to obtain a nonpositive term prior to weak death. But, for all we know, the sequence may strongly die, persist, or live. We believe that it lives.

The rest of this chapter consists of five sections that provide some background on nested recurrences. Section 2.1 summarizes some of the notation we use to refer to nested recurrences and their solutions. Then, Sections 2.2 and 2.3 describe the sorts of nested-recurrent sequences that have been previously analyzed. Most of our results in the forthcoming chapters fall into these categories as well. Section 2.4 defines the notation of a linear nested recurrence (the primary type we consider), and it gives a self-contained proof of the shift-invariance of solutions to certain linear nested recurrences. Finally, Section 2.5 contains a preview of the upcoming chapters and how they tie in with the background given in this chapter.

## 2.1 Notation

As mentioned earlier, we use the notation $\langle a_1, a_2, \ldots, a_k \rangle$ to denote an initial condition of length $k$. Throughout this dissertation, $Q(n)$ is used to denote the $n^{th}$ term of the Hofstadter $Q$-sequence, and $F(n)$ is used to denote the $n^{th}$ Fibonacci number. We frequently consider the Hofstadter $Q$-recurrence with different initial conditions; any other sequence satisfying the $Q$-recurrence is denoted by $Q_\alpha$ for some subscript $\alpha$. The notation for a given example is introduced when the sequence is introduced, and the subscript generally has some connection to the sequence. (For instance, Golomb's solution in the next section is denoted by $Q_G$.) Other nested recurrences are given symbols other than $Q$ or $F$ to represent them. If the recurrence has standard notation, we use that, and, otherwise, we assign it an unused letter. If the same recurrence subsequently appears with a different initial condition, the same letter with subscripts will be used, as with the $Q$-recurrence.

## 2.2 Linear-Recurrent Solutions

Golomb [15] considers the result of replacing the initial condition of the $Q$-sequence with the new initial condition $\langle 3, 2, 1\rangle$. The first few terms of the resulting sequence are $3, 2, 1, 3, 5, 4, 3, 8, 7, 3, 11, 10, \ldots$. This sequence lives, and the pattern visible here continues forever [31, A244477]. The result is the linear-recurrent sequence defined by

$$\begin{cases} Q_G(3k) = 3k - 2 \\ Q_G(3k+1) = 3 \\ Q_G(3k+2) = 3k+2. \end{cases} \tag{2.1}$$

Golomb's sequence is an example of a *quasilinear* sequence, that is, a sequence consisting of interleaved polynomial sequences of degree at most one. In general, a *quasipolynomial* of degree $d$ is a sequence consisting of interleaved polynomial sequences of degree at most $d$.

We give a proof here of Golomb's result, not because the proof is particularly compelling or difficult, but because this proof and its structure provide the underpinning for much of the work of this dissertation.

*Proof.* The proof is by induction on $k$. A typical inductive proof begins with a base case. But, for illustrative purposes, we perform the inductive step first and the base case last.

Suppose that (2.1) holds for all $k' < k$. There are three cases to consider:

**Index $3k$:** We have

$$\begin{aligned} Q_G(3k) &= Q_G(3k - Q_G(3k-1)) + Q_G(3k - Q_G(3k-2)) \\ &= Q_G(3k - (3(k-1)+2)) + Q_G(3k-3) \\ &= Q_G(1) + Q_G(3(k-1)) \\ &= 3 + 3(k-1) - 2 \\ &= 3k-2, \end{aligned}$$

as required.

**Index $3k+1$:** We have

$$\begin{aligned} Q_G(3k+1) &= Q_G(3k+1-Q_G(3k)) + Q_G(3k+1-Q_G(3k-1)) \\ &= Q_G(3k+1-(3k-2)) + Q_G(3k+1-(3\,(k-1)+2)) \\ &= Q_G(3) + Q_G(2) \\ &= 3, \end{aligned}$$

as required.

**Index $3k+2$:** We have

$$\begin{aligned} Q_G(3k) &= Q_G(3k+2-Q_G(3k+1)) + Q_G(3k+2-Q_G(3k)) \\ &= Q_G(3k+2-3) + Q_G(3k+2-(3k-2)) \\ &= Q_G(3k-1) + Q_G(4) \\ &= 3k-1+3 \\ &= 3k+2, \end{aligned}$$

as required.

We now return to the base case. We must answer the question: "What do we assume in the inductive step?" For one thing, we assume that the values of $Q_G(1)$, $Q_G(2)$, $Q_G(3)$, and $Q_G(4)$ are 3, 2, 1, and 3, respectively. We also assume that $k \geq 1$, as we refer to earlier terms of (2.1) with $k' = k-1$. We only refer to $Q_G(4)$ in the $3k+2$ case, so the inductive step actually succeeds in calculating $Q_G(4)$ (as 4 is congruent to 1 mod 3). As a result, our base case needs to be the first three values. These are precisely the initial conditions, and they satisfy (2.1). □

Ruskey [30] instead considers the initial condition $\langle 3, 6, 5, 3, 6, 8 \rangle$ for the $Q$-recurrence. The resulting sequence weakly dies immediately (as $Q_R(7)$ depends on $Q_R(-1)$). But, it persists, and every third term is a Fibonacci number [31, A188670]:

$$\begin{cases} Q_R(3k) = F(k+4) \\ Q_R(3k+1) = 3 \\ Q_R(3k+2) = 6. \end{cases} \tag{2.2}$$

The proof of (2.2) is quite similar to the proof of (2.1).

The solutions of Golomb and Ruskey both satisfy linear recurrences ($a_n = 2a_{n-3} - a_{n-6}$ in the case of Golomb; $a_n = 2a_{n-3} - a_{n-9}$ in the case of Ruskey). One might hope for a method that, given a linear recurrence and a Hofstadter-like recurrence, determines whether there is a sequence that eventually satisfies both of them. This is quite a lofty goal. But, we see that Golomb's sequence and Ruskey's sequence both have a deeper structure. Golomb's sequence is a quasipolynomial with period 3. Ruskey's sequence, while not a quasipolynomial, is also structured as an interleaving of three simpler sequences. Going forward, this is the key feature of such solutions that we fixate on.

## 2.3 Slow Solutions

Most of the recent literature on nested recurrences has been devoted to the study of so-called *slow solutions* [9, 20–23, 32].

**Definition 2.4.** *A sequence of integers* $(a_n)$ *is called* slow *if* $a_n - a_{n-1} \in \{0, 1\}$ *for all indices* $n$.

In contrast to the solutions from the previous section, slow sequences often arise from varying the recurrence while retaining the initial condition of the $Q$-sequence. The first notable example of a slow solution to a nested recurrence is due to Conolly [7]. If the recurrence $C(n) = C(n - C(n-1)) + C(n-1-C(n-2))$ is given the initial condition $C(1) = C(2) = 1$, the result is a slow sequence (see Figure 2.2) where every $k > 1$ appears once for each factor of 2 that divides it [31, A046699]. As a consequence of this simple description, the sequence satisfies

$$\lim_{n\to\infty} \frac{C(n)}{n} = \frac{1}{2}.$$

Notably, the only difference between the Conolly recurrence and the Hofstadter $Q$-recurrence is the "$-1$" in the second term. In this way, the Conolly recurrence is, in some sense, a "shifted" version of the Hofstadter recurrence.

Many other examples of slow sequences are known [32]. Many of these use recurrences with shifts similar to the one in Conolly's recurrence. But, perhaps the most

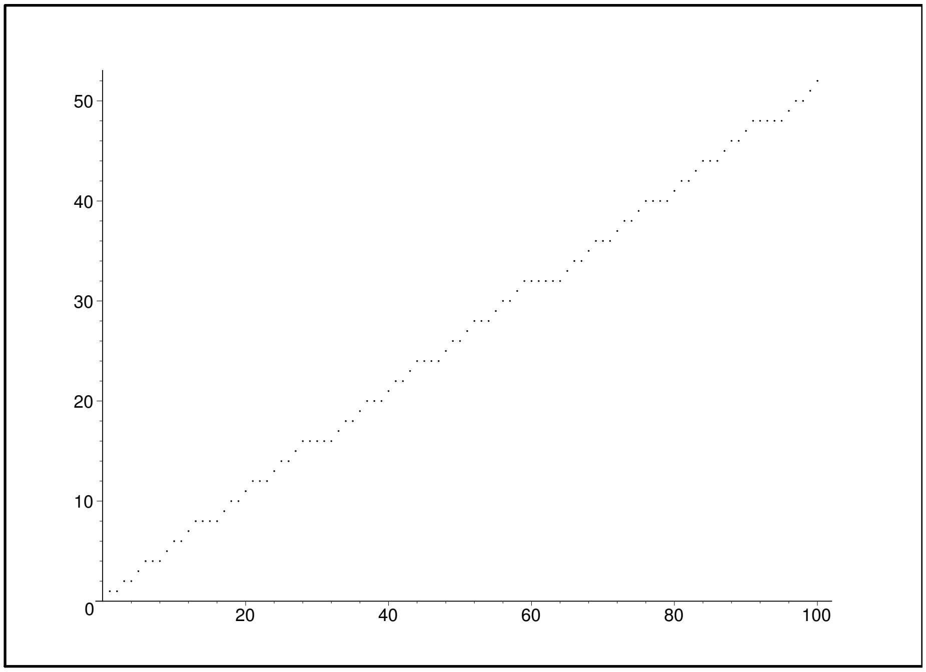


Figure 2.2: The first 100 terms of Conolly's sequence (A046699 in OEIS)

famous example of a slow solution to a nested recurrence is the Hofstadter-Conway \$10000 Sequence [31, A004001], given by $A(n) = A(A(n-1)) + A(n - A(n-1))$ with $A(1) = A(2) = 1$. (We call a term like the $A(A(n-1))$ in this recurrence a *Conway-like term.*) John H. Conway was able to show that

$$\lim_{n\to\infty} \frac{A(n)}{n} = \frac{1}{2},$$

but he had no clue about the rate of convergence. (See Figure 2.3.) Notably, Conway offered a \$10000 prize for an analysis of the behavior. Colin Mallows solved this problem a few years later [28], though Conway, at that point, tried to claim to have offered only \$1000.

Hofstadter and Huber [3, 19] have investigated the following family of recurrences, which generalize the Hofstadter $Q$-recurrence. For integers $0 < r < s$, define

$$Q_{r,s}(n) = Q_{r,s}(n - Q_{r,s}(n-r)) + Q_{r,s}(n - Q_{r,s}(n-s)).$$

They explore these recurrences experimentally for various initial conditions. This work has led them to conjecture that the sequences resulting from an all-ones initial condition always die, except for $(r, s) \in \{(1, 1), (1, 4), (2, 4)\}$. The case $(1, 1)$ is the $Q$-sequence,

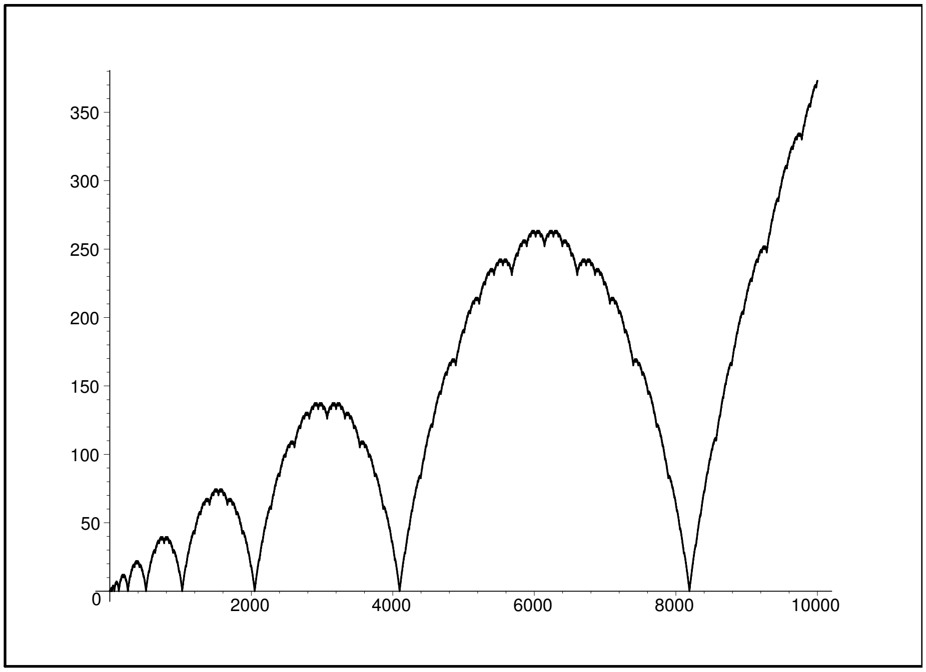


Figure 2.3: First 10000 values of $A(n) - \frac{n}{2}$ (Hofstadter-Conway)

and the case $(2, 4)$, often called the $W$-sequence [31, A087777] (see Figure 2.4), displays even wilder behavior than the $Q$-sequence. On the other hand, the sequence resulting from $(r, s) = (1, 4)$ (see Figure 2.5), known as the $V$-sequence [31, A063882], is slow [3]. In fact, this sequence was recently shown to be 2-automatic [1]. This slow sequence results from a recurrence with no Conolly-like shifts and with no Conway-like terms.

There are two methods commonly used to prove that Hofstadter-like sequences are slow. In some cases, there are combinatorial interpretations for slow sequences involving counting leaves in nested tree structures [22]. The sequence counting the leaves is obviously slow; the main difficulty comes in showing that the nested recurrence also describes the same structure. For some slow sequences, though, there is no known combinatorial interpretation. The other proofs of slowness usually go by induction with complicated inductive hypotheses. For a sequence $(a_n)$, one would love to work with just the inductive hypothesis that $a_m - a_{m-1} \in \{0, 1\}$ for all $m < n$, but this is never enough. Instead, additional inductive hypotheses are required to handle certain cases. These extra hypotheses strongly depend on the sequence in question. While the sequences like Conolly's, resulting from "shifted" recurrences, have similar proofs to each other, the

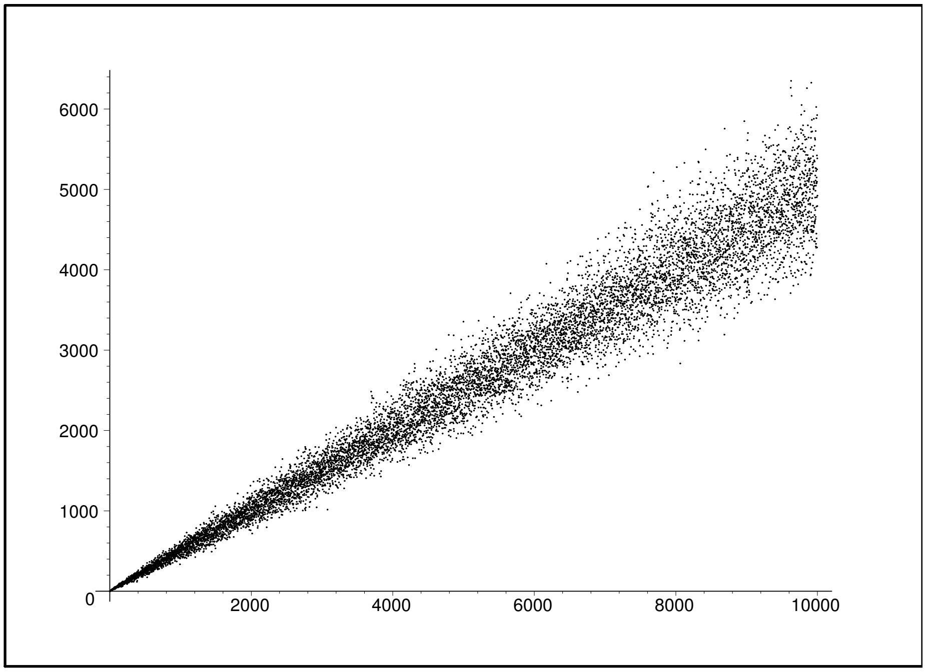


Figure 2.4: First 10000 values of $W(n)$ (A087777 in OEIS)

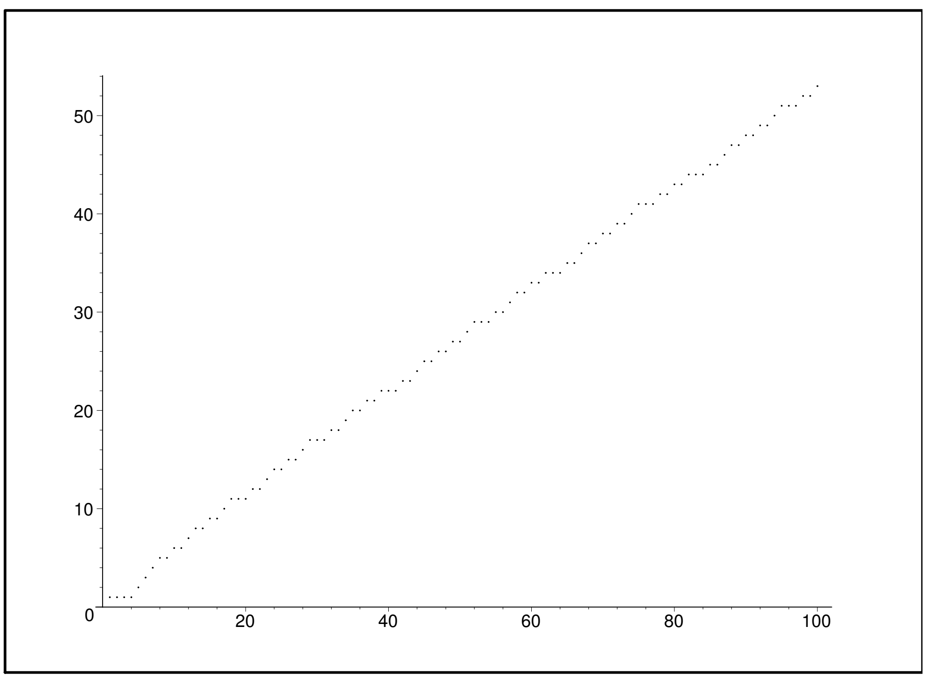


Figure 2.5: First 100 values of $V(n)$ (A063882 in OEIS)

proof for the $V$-sequence uses a different sort of inductive hypothesis. One would like to automate these proofs, but this would require some method of deciding the appropriate inductive hypotheses.

## 2.4 Linear Nested Recurrences and Shift Invariance

The primary objects of study we study are *linear nested recurrences.*

**Definition 2.5.** *A nested recurrence $L$ is* linear *if it is of the form*

$$L(n) = P(n) + \sum_{i=1}^{d} \alpha_i L(E_i),$$

*where $P(n)$ is an explicit expression in $n$, $d$ is a nonnegative integer, each $\alpha_i$ is an integer, and each $E_i$ is an expression of the same form as the generic formula for $L(n)$ (thereby allowing for arbitrarily many nesting levels).*

Going forward, when we refer to a recurrence as *linear*, we mean linear nested if the recurrence is nested.

The form of linear nested recurrences is quite general. Most commonly, we have the following:

**Definition 2.6.** *In a linear nested recurrence $L(n)$, if*

- *$P(n)$ is a polynomial,*
- *Each $\alpha_i$ is positive,*
- *Each $E_i$ of the form $n - \beta_i - \sum_{j=1}^{d_i} L(n - \gamma_{i,j})$,*
- *Each $d_i$ is a positive integer,*
- *Each $\beta_i$ is a nonnegative integer,*
- *Each $\gamma_{i,j}$ is a positive integer,*

*then $L$ is* basic linear *(or just* basic*).*

Most commonly, we actually have $P(n) = 0$ and each $d_i = 1$. The Hofstadter $Q$-recurrence is basic, as $P(n) = 0$, $d = 2$, $d_1 = 1$, $d_2 = 1$, $\alpha_1 = \alpha_2 = 1$, $E_1 = n - Q(n-1)$, and $E_2 = n - Q(n-2)$. On the other hand, the Hofstadter-Conway recurrence, while linear, is not basic, because of the term $A(A(n-1))$.

A key property of basic linear recurrences is the so-called *shift invariance* of their solutions.

**Proposition 2.7.** *Let $L(n)$ be a basic linear nested recurrence with $P(n) = c$ for some integer $c$. Suppose $(a_n)_{1\leq n\leq N}$ (where $1 \leq N \leq \infty$) is a solution to $L$ under the strong death convention, with the initial condition $\langle a_1, a_2, \ldots, a_k\rangle$. Define a sequence $(b_n)_{1\leq n\leq N+1}$ by*

$$\begin{cases} b_1 = 0 \\ b_n = a_{n-1} \quad 2 \leq n \leq N+1. \end{cases}$$

*Then, $(b_n)$ is also a solution to $L$, and it is generated by $\langle b_1, b_2, \ldots, b_{k+1}\rangle$.*

*Proof.* Since we are using the strong death convention, we may extend $(a_n)$ and $(b_n)$ to nonpositive indices by defining all such terms to be zero. So, we actually have $b_n = a_{n-1}$ for all $n$.

We have that

$$L(n) = c + \sum_{i=1}^{d} \alpha_i L\left(n - \beta_i - \sum_{j=1}^{d_i} L(n - \gamma_{i,j})\right),$$

and we must show for all $n > k+1$ that $b_n = L(n)$ when $\langle b_1, b_2, \ldots, b_{k+1}\rangle$ is given as the initial condition. We do so by induction.

Suppose that $n > k+1$, and furthermore suppose that $b_m$ satisfies $L$ for every $k+1 < m < n$. Since $(a_n)$ is a solution to $L$, we sometimes write $L'(n)$ to denote $a_n$. We have

$$\begin{aligned} L(n) &= c + \sum_{i=1}^{d} \alpha_i L\left(n - \beta_i - \sum_{j=1}^{d_i} L(n - \gamma_{i,j})\right) \\ &= c + \sum_{i=1}^{d} \alpha_i b_{n - \beta_i - \sum_{j=1}^{d_i} L(n-\gamma_{i,j})} \end{aligned}$$

$$= c + \sum_{i=1}^{d} \alpha_i b_{n-\beta_i - \sum_{j=1}^{d_i} b_{n-\gamma_{i,j}}}$$

$$= c + \sum_{i=1}^{d} \alpha_i a_{n-1-\beta_i - \sum_{j=1}^{d_i} a_{n-1-\gamma_{i,j}}}$$

$$= c + \sum_{i=1}^{d} \alpha_i L'\left(n - 1 - \beta_i - \sum_{j=1}^{d_i} L'(n-1-\gamma_{i,j})\right)$$

$$= L'(n-1) = a_{n-1} = b_n,$$

as required. □

In the forthcoming chapters, we frequently make use of shift invariance.

## 2.5 Preview of Upcoming Chapters

The remainder of this dissertation is structured as follows. Chapter 3 describes an algorithm for automatically discovering, with proof, linear-recurrent solutions to nested recurrences (in the vein of Golomb's and Ruskey's sequences). Then, Chapters 4 and 5 describe two results that were originally obtained via explorations using the algorithm in Chapter 3. Chapter 6, which follows, is a fairly self-contained description of a new slow sequence, which superficially resembles the $V$-sequence in a few ways. Then, Chapter 7 introduces a system of three nested recurrences. The resulting sequences are enigmatic on their own, and together they are interwoven into certain solutions to the $Q$-recurrence. The next three chapters, Chapter 8, Chapter 9, and Chapter 10 initiate a novel methodology for studying nested recurrent sequences. Whereas previous studies have been primarily concerned with finding initial conditions that lead to prescribed behavior, these chapters prescribe the form of the initial condition and determine the resulting sequence behaviors using computational methods. Finally, Chapter 11 summarizes some open problems about nested recurrences and suggests some future research directions.

Various Maple programs implementing all of the algorithms in this dissertation, as well as some related procedures, can be found at `http://github.com/nhf216/thesis`.

Generally speaking, the procedures in this package offer more general versions of the algorithms described in this dissertation.

# Chapter 3

# Discovering Nice Solutions to Nested Recurrences Using Symbolic Computation

In the previous chapter, in Section 2.2, we introduced the sequences $Q_G$ and $Q_R$, Golomb's [15] and Ruskey's [30] respective solutions to the Hofstadter $Q$-recurrence. In this chapter, we develop a machine to find similar solutions (under the notion of weak death). In Section 3.1, we introduce a formalism that encapsulates the notion of an interleaving of simple sequences. Then, in Section 3.2, we describe our algorithm that finds these special solutions. Finally, in Section 3.3 we describe some notable sequences found using the methods in this chapter.

A Maple package implementing all of the algorithms in this chapter, as well as some related procedures, can be found at `http://github.com/nhf216/thesis/nicehof.txt`.

## 3.1 Positive Recurrence Systems

We observed previously that both Golomb's and Ruskey's solutions are interleavings of three simple sequences. But, in order to generalize these, we need a rigorous notion of what a "simple sequence" is. In the case of Ruskey's solution, each of the three interleaved sequences can be described by a homogeneous linear recurrence with nonnegative coefficients:

$$\begin{cases} a_k = a_{k-1} & a_0 = 3 \\ b_k = b_{k-1} & b_0 = 6 \\ c_k = c_{k-1} + c_{k-2} & c_1 = 5, c_2 = 8. \end{cases}$$

(Note that these recurrences are not unique.) Golomb's solution cannot be expressed in this way, but each of its interleaved sequences can be described by a nonhomogeneous linear recurrence with nonnegative coefficients:

$$\begin{cases} a_k = 3 + a_{k-1} & a_0 = 3 \\ b_k = 3 & \\ c_k = 3 + c_{k-1} & c_1 = 1. \end{cases}$$

In both of these cases, we have a system of three nonhomogneous linear recurrences where all coefficients are nonnegative. Here, none of the recurrences refer to each other in their definitions, but we want to allow for this possibility. This leads to the following generalization:

**Definition 3.1.** *A* positive recurrence system *is a system of $m$ nonhomogeneous linear recurrences of the form*

$$\begin{cases} a_k^{(1)} = P_1(k) + \sum_{\ell=1}^{d} \sum_{j=1}^{m} \alpha_{1,\ell,j} a_{k-\ell}^{(j)} \\ a_k^{(2)} = P_2(k) + \sum_{\ell=1}^{d} \sum_{j=1}^{m} \alpha_{2,\ell,j} a_{k-\ell}^{(j)} \\ \vdots \\ a_k^{(m)} = P_m(k) + \sum_{\ell=1}^{d} \sum_{j=1}^{m} \alpha_{m,\ell,j} a_{k-\ell}^{(j)} \end{cases}$$

*satisfying the following conditions:*

- *$d$ is a nonnegative integer.*
- *$P_1(k)$ through $P_m(k)$ are integer-valued polynomials that take nonnegative values for sufficiently large $k$. (We call such polynomials* eventually nonnegative.*)*
- *Each $\alpha_{i,\ell,j}$ is a nonnegative integer.*

Note that, for convenience, we may sometimes have a recurrence system where $a_k^{(i)}$ refers to $a_k^{(j)}$ for some $j < i$. This is permissible, as we can just replace $a_k^{(j)}$ with its right-hand side in order to conform to Definition 3.1.

The solutions to Hofstadter-like recurrences that we seek are eventual interleavings of sequences that together satisfy a positive recurrence system. What follows is a formalization of this notion:

**Definition 3.2.** *Let $m$ be a positive integer. The sequence $(a_k)_{k\geq 1}$ is* positive-recurrent *with period $m$ if there exists an integer $K$ such that the sequences*

$$\left\{(a_{ms+r})_{s\geq K} : 0 \leq r < m\right\}$$

*satisfy a positive recurrence system.*

Observe that eventually nonnegative polynomials are trivially positive-recurrent with period 1, as we can take $d = 0$. Also, any sequence satisfying a homogeneous linear recurrence with nonnegative coefficients is positive-recurrent with period 1, as we can take $P_1$ identically 0.

Any positive-recurrent sequence is itself eventually linear recurrent, as are all of the component sequences. This is true because a positive recurrence system can be converted into a linear system of equations for the generating functions of the component sequences. Each generating function is therefore a rational function. The resulting linear recurrence may have some negative coefficients, though. The ability to convert a generic linear recurrent sequence to a positive-recurrent sequence is less clear, even if the values are eventually positive. The potential difficulty comes entirely from the positivity requirements on the coefficients in a positive recurrence system, along with the permissibility of nonhomogeneous parts in such systems.

We are concerned with determining the rate of growth of each sequence in a solution to a positive recurrence system. In order for things to be well-defined and easy to analyze, we will need the following technical definition.

**Definition 3.3.** *An initial condition of length $N$ to a positive recurrence system is called* eventually positive *if the following conditions hold:*

- *If $k \geq N$, then $P_r(k) \geq 0$ for all $r$ and all $k$.*
- *For all $0 \leq i \leq d$, $a^{(r)}_{N-i} > 0$ for all $r$.*

Any long enough positive initial condition is eventually positive. But, this definition allows for some nonpositive values early in the initial condition, so long as those values are never used in calculating recursively defined terms. Furthermore, we require all of the polynomials to be nonnegative when calculating recursive terms. This will be useful in our analysis, though it is not strictly necessary. (A much more complicated, weaker condition would suffice, and, in that case, we would not even need all of the polynomials to be eventually nonnegative.)

In the case where we have a solution to a positive recurrence system given by an eventually positive initial condition, the following algorithm determines the order of growth of each component sequence. This algorithm is based on graphs; for definitions of any unfamiliar terms, see, for example, Chapter 1 of [4]. In particular, all paths, cycles, and circuits we consider are directed.

1. Define a weighted directed graph $G$ as follows:

   - The vertices of $G$ are the integers $\{1, \ldots, m\}$.
   - There is an edge from $i$ to $j$ if and only if, for some $\ell$, $\alpha_{i,\ell,j} > 0$.
   - The weight of the edge from $i$ to $j$ is

   $$\sum_{\ell=1}^{d} \alpha_{i,\ell,j}.$$

2. Initialize variables $d_1, d_2, \ldots, d_m$ so that $d_i$ equals the degree of $P_i$.

3. Let $W$ denote the set of vertices $v$ in $G$ satisfying one of the following:

   - $v$ is in a circuit with at least one edge having weight greater than 1.
   - $v$ is in more than one circuit.

   For each $v \in W$, set $d_v$ to $\infty$ and delete any outgoing edge from $v$ in $G$ that is part of a cycle. Call the resulting graph $G'$. (We can actually delete *all* outgoing edges from $v$, but the form we have stated here will be more useful when we prove this algorithm's correctness.)

4. Define the following relation $\sim$ on $\{1, 2, \ldots, m\}$:

   $$i \sim j \text{ if and only if } (i = j) \text{ or } (i \text{ and } j \text{ are in a cycle together in } G').$$

   As a consequence of Step 3, it is easy to check that $\sim$ is an equivalence relation. Each equivalence class is either a single vertex or a cycle.

5. Define a directed graph $H$ as follows:

   - The vertices of $H$ are the equivalence classes of $\{1, 2, \ldots, m\}$ under $\sim$.
   - There is an edge from class $I$ to class $J$ if and only if there is an edge in $G'$ from some $i \in I$ to some $j \in J$.

   If $H$ contains a cycle $I_1, I_2, \ldots, I_q$, then for each $1 \leq h < q$, there is an edge in $G'$ from some $i_h^{out} \in I_h$ to some $i_{h+1}^{in} \in I_{h+1}$. Also, there is an edge in $G'$ from some $i_q^{out} \in I_q$ to some $i_1^{in} \in I_1$. Furthermore, by the definition of $\sim$, for each $h$, there is a (possibly trivial) path from $i_h^{in}$ to $i_h^{out}$ within $I_h$. Concatenating all of these edges together gives a cycle in $G'$ that includes elements of multiple equivalence classes, which contradicts the definition of $G'$.

   So, we can conclude that $H$ contains no cycles.

6. For each vertex $I$ of $H$, initialize a variable $d_I = \max_{i \in I} d_i$.

7. Topologically sort the vertices of $H$. Consider the vertices $I$ from last to first:

   - If $I$ is a cycle in $G'$ (including a single vertex with a loop), set $d_I$ to $d_I + 1$ (unless $d_I$ was $-\infty$, in which case it should be set to 0).
   - For all $J$ with an edge from $J$ to $I$, set $d_J = \max(d_J, d_I)$.

At the end of this process, we have values $d_I$ for each equivalence class $I$. In general, for an integer $r$, let $\bar{r}$ denote its equivalence class under $\sim$. We now make the following claim:

**Claim 3.4.** *Suppose we have a positive recurrence system with $m$ component sequences, along with an eventually positive initial condition. Let $1 \leq r \leq m$ be an integer. If $d_{\bar{r}} < \infty$, then $a_k^{(r)} = \Theta\left(k^{d_{\bar{r}}}\right)$. If $d_{\bar{r}} = \infty$, then $a_k^{(r)}$ grows exponentially.*

*Proof.* Since each component sequence has a rational generating function, no component sequence can possibly grow faster than exponentially. So, proving only lower bounds when sequences should grow exponentially suffices. (We take advantage of this simplification throughout this proof.)

For each vertex $r$ in a directed graph $F$, define the *potential* of $r$ (denoted $\phi_F(r)$) as the sum of the lengths of all directed cycles in $F$ containing $r$ (or 0 if $r$ is not in any cycles). Also, for each $r \in \langle m \rangle$, define $J(r)$ as the set of immediate successors of $r$ in $G'$. We will prove Claim 3.4 by induction on $\phi_G(r)$. We examine three cases (essentially a special case, a base case, and an inductive step).

**$\phi_G(r) = 0$:** In this case, $r$ is in no cycles in $G$, and $\bar{r} = \{r\}$. Also, every successor of $r$ in $G$ is in $J(r)$, and every coefficient $\alpha_{r,\ell,j}$ with $j \notin J(r)$ is zero. Inductively, suppose that any $j \in J(r)$ satisfies Claim 3.4. (The base case of $J(r) = \emptyset$ is implicitly included here.) We have

$$a_k^{(r)} = P_r(k) + \sum_{\ell=1}^{d} \sum_{j \in J(r)} \alpha_{r,\ell,j} a_{k-\ell}^{(j)}.$$

Let $D = \max\left\{d_{\bar{j}} : j \in J(r)\right\}$ (or $D = -\infty$ if this set is empty). If $D = \infty$, then, by induction, some sequence $a_k^{(j)}$ with $j \in J(r)$ grows exponentially. As a result, $a_k^{(r)}$ grows exponentially, and we have $d_{\bar{r}} = \infty$, as required. On the other hand, if $D < \infty$, then we have $a_k^{(r)} = \Theta\left(k^{\max(D, \deg(P_r))}\right)$. Since $d_{\bar{r}} = \max(D, \deg(P_r))$ in this case, we have $a_k^{(r)} = \Theta\left(k^{d_{\bar{r}}}\right)$, as required.

**$\phi_G(r) = 1$:** In this case, the only cycle of $G$ that $r$ is in is a self-loop, and $\bar{r} = \{r\}$. Also, every non-$r$ successor of $r$ in $G$ is in $J(r)$, and every coefficient $\alpha_{r,\ell,j}$ with $j \notin \{r\} \cup J(r)$ is zero. Inductively, suppose that any $j \in J(r)$ satisfies Claim 3.4. (The base case of $J(r) = \emptyset$ is implicitly included here.) We have,

$$a_k^{(r)} = P_r(k) + \sum_{\ell=1}^{d} \alpha_{r,\ell,r} a_{k-\ell}^{(r)} + \sum_{\ell=1}^{d} \sum_{j \in J(r)} \alpha_{r,\ell,j} a_{k-\ell}^{(j)}.$$

Let $D = \max\left\{d_{\bar{j}} : j \in J(r)\right\}$ (or $D = -\infty$ if this set is empty). If $D = \infty$, then, by induction, some sequence $a_k^{(j)}$ with $j \in J(r)$ grows exponentially. As a result, $a_k^{(r)}$ grows exponentially, and we have $d_{\bar{r}} = \infty$, as required. On the other hand, if $D < \infty$, then we have

$$a_k^{(r)} = \sum_{\ell=1}^{d} \alpha_{r,\ell,r} a_{k-\ell}^{(r)} + \Theta\left(k^{\max(D, \deg(P_r))}\right).$$

This is a nonhomogeneous linear recurrence for $a_k^{(r)}$ with characteristic polynomial

$$P(x) = x^d - \sum_{\ell=1}^{d} \alpha_{r,\ell,r} x^{d-\ell}$$

If the weight on the self-loop on $r$ is 1, then

$$\sum_{\ell=1}^{d} \alpha_{r,\ell,r} = 1.$$

So, $p(1) = 0$, and we also note that this is a simple root, and the other nonzero roots of $p$ are all simple and roots of unity. By the theory of linear recurrences, this results in a solution where $a_k^{(r)} = \Theta\left(k^{1+\max(D,\deg(P_r))}\right)$ (unless $D$ and $\deg(P_r)$ are both $-\infty$, in which case $a_k^{(r)} = \Theta(1)$). Since $d_{\bar{r}} = 1 + \max(D, \deg(P_r))$ in this case, we have $a_k^{(r)} = \Theta\left(k^{d_{\bar{r}}}\right)$, as required.

If the weight on the self-loop on $r$ is greater than 1, then

$$\sum_{\ell=1}^{d} \alpha_{r,\ell,r} > 1.$$

So, $p(1) < 0$, but $p(x) \to \infty$ as $x \to \infty$. So, $p$ has a real root that is greater than 1. By the theory of linear recurrences, we have a solution where $a_k^{(r)}$ grows exponentially. And, $d_{\bar{r}} = \infty$ here, as required.

**$\phi_G(r) > 1$:** Here $r$ is in some cycle of length at least 2. Consider such a cycle $C$. Let $s$ denote the immediate successor of $r$, and let $t$ denote the immediate successor of $s$. We now construct a new positive recurrence system, and if we run our algorithm on this new system, we will obtain a graph in step 1. Call this graph $\tilde{G}$. The system and $\tilde{G}$ should have the following properties:

- The system has $m-1$ component sequences

$$b_k^{(1)}, b_k^{(2)}, \ldots, b_k^{(s-2)}, b_k^{(s-1)}, b_k^{(s+1)}, b_k^{(s+2)}, \ldots, b_k^{(m)},$$

  one corresponding to each sequence $a_k^{(i)}$ with $i \neq s$.

- For every $i \neq s$, if $b_k^{(i)}$ is given the same initial condition as $a_k^{(i)}$, then $b_k^{(i)} \leq a_k^{(i)}$.

- If $s$ is in a unique cycle in $G$, then $b_k^{(r)} = a_k^{(r)}$.
- We have $\phi_{\tilde{G}}(r) < \phi_G(r)$.

Hence, if we inductively assume that Claim 3.4 holds for any component sequence $r'$ of smaller potential in *any* positive recurrence system, this construction completes the proof. (The properties are sufficient to show that $a_k^{(r)}$ grows at least as fast as it should, and the equality case handles the sub-exponential component sequences.)

We have

$$\begin{aligned} a_k^{(s)} &= P_s(k) + \sum_{\ell=1}^{d}\sum_{j=1}^{m} \alpha_{s,\ell,j} a_{k-\ell}^{(j)} \\ &\geq P_s(k) + \sum_{\ell=1}^{d}\left(\alpha_{s,\ell,t} a_{k-\ell}^{(t)} + \sum_{j\in J(s)} \alpha_{s,\ell,j} a_{k-\ell}^{(j)}\right). \end{aligned}$$

If $s$ is in a unique cycle in $G$, then every immediate successor of $s$ other than $t$ is in $J(s)$. This means that all dropped terms in going from the equality to the inequality are in fact zero, so we actually have equality here in this case.

We also have

$$\begin{aligned} a_k^{(r)} &= P_r(k) + \sum_{\ell=1}^{d}\sum_{j=1}^{m} \alpha_{r,\ell,j} a_{k-\ell}^{(j)} \\ &= P_r(k) + \sum_{\ell=1}^{d}\left(\alpha_{r,\ell,s} a_{k-\ell}^{(s)} + \sum_{j\in\langle m\rangle\setminus\{s\}} \alpha_{r,\ell,j} a_{k-\ell}^{(j)}\right). \end{aligned}$$

Let $J' = \langle m\rangle \setminus (\{t\} \cup J(s))$. Combining everything together yields

$$\begin{aligned} a_k^{(r)} &\geq P_r(k) + \sum_{\ell_1=1}^{d}\left(\alpha_{r,\ell_1,s}\left(P_s(k-\ell_1) + \sum_{\ell_2=1}^{d}\left(\alpha_{s,\ell_2,t} a_{k-\ell_1-\ell_2}^{(t)}\right.\right.\right. \\ &\qquad \left.\left.\left. + \sum_{j\in J(s)} \alpha_{s,\ell_2,j} a_{k-\ell_1-\ell_2}^{(j)}\right)\right) + \sum_{j\in\langle m\rangle\setminus\{s\}} \alpha_{r,\ell_1,j} a_{k-\ell_1}^{(j)}\right) \\ &= \left(P_r(k) + \sum_{\ell=1}^{d} \alpha_{r,\ell,s} P_s(k-\ell)\right) \\ &\qquad + \sum_{\ell_1=1}^{d}\left(\alpha_{r,\ell_1,t} a_{k-\ell_1}^{(t)} + \sum_{\ell_2=1}^{d} \alpha_{s,\ell_2,t} a_{k-\ell_1-\ell_2}^{(t)}\right) \end{aligned}$$

$$+\sum_{\ell_1=1}^{d}\sum_{j\in J(s)}\left(\alpha_{r,\ell_1,j}a^{(j)}_{k-\ell_1}+\sum_{\ell_2=1}^{d}\alpha_{s,\ell_2,j}a^{(j)}_{k-\ell_1-\ell_2}\right)$$
$$+\sum_{\ell=1}^{d}\sum_{j\in J'}\alpha_{r,\ell,j}a^{(j)}_{k-\ell}.$$

Again, if $s$ is in a unique cycle in $G$, we actually have equality here.

Our new positive recurrence system is obtained from the current one as follows:

1. The equation for $b_k^{(i)}$ for $i \notin \{r, s\}$ is

$$b_k^{(i)} = P_i(k) + \sum_{\ell=1}^{d}\sum_{j\in\langle m\rangle\setminus\{s\}}\alpha_{i,\ell,j}b^{(j)}_{k-\ell}.$$

2. We have

$$b_k^{(r)} = \left(P_r(k)+\sum_{\ell=1}^{d}\alpha_{r,\ell,s}P_s(k-\ell)\right)$$
$$+\sum_{\ell_1=1}^{d}\left(\alpha_{r,\ell_1,t}b^{(t)}_{k-\ell_1}+\sum_{\ell_2=1}^{d}\alpha_{s,\ell_2,t}b^{(t)}_{k-\ell_1-\ell_2}\right)$$
$$+\sum_{\ell_1=1}^{d}\sum_{j\in J(s)}\left(\alpha_{r,\ell_1,j}b^{(j)}_{k-\ell_1}+\sum_{\ell_2=1}^{d}\alpha_{s,\ell_2,j}b^{(j)}_{k-\ell_1-\ell_2}\right)$$
$$+\sum_{\ell=1}^{d}\sum_{j\in J'}\alpha_{r,\ell,j}b^{(j)}_{k-\ell}.$$

In the definition of $b_k^{(r)}$, the first part is the nonhomogeneous part: a polynomial of degree $\max\left(\deg\left(P_r\right),\deg\left(P_s\right)\right)$. The second part describes the references to $\left(b_k^{(t)}\right)$. The third part describes the references to $\left(b_k^{(j)}\right)$ where $j \in J(s)$. The final part describes all the other references. (Note that there are no references to the nonexistent sequence $\left(b_k^{(s)}\right)$.)

From this construction and our calculation, we notice that the first three desired properties of this new system definitely hold. We can also see how $\tilde{G}$ is obtained from $G$:

- First, delete all incoming edges to $s$, other than the one from $r$. (This is accomplished by item 1 above.)

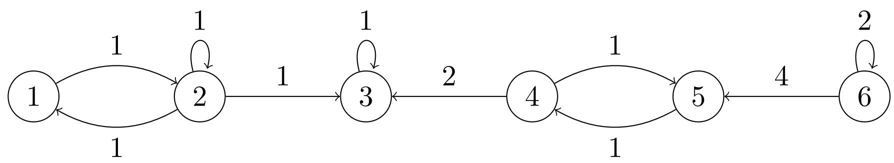


Figure 3.1: The graph $G$ in our example

- Then, contract the edge $rs$. (This is accomplished by item 2 above.)
- If any multiple edges were created in the previous step, replace them by a single edge whose weight is the sum of the weights of the edges being replaced. (This would come up, for example, if $r$ has a self-loop and is also in a 2-cycle, with $s$ being the other vertex in that cycle. After contracting $rs$, $r$ would have two self-loops, which then need to be combined.)

It is clear that every cycle in $G$ containing $r$ corresponds to a cycle in $\tilde{G}$ containing $r$, and the latter cycle can be no longer than the original. The cycle corresponding to $C$ in $\tilde{G}$ is strictly shorter than $C$, since $s$ has been removed from it. So, $\phi_{\tilde{G}}(r) < \phi_G(r)$, as required.

□

We conclude this section with an example run of this graph algorithm. We shall determine the asymptotic behavior of an arbitrary eventually positive solution to the following positive recurrence system:

$$\begin{cases} a_k^{(1)} = a_{k-1}^{(2)} \\ a_k^{(2)} = a_{k-1}^{(1)} + a_{k-2}^{(2)} + a_{k-1}^{(3)} \\ a_k^{(3)} = k^2 + a_{k-1}^{(3)} \\ a_k^{(4)} = 1 + a_{k-1}^{(3)} + a_{k-2}^{(3)} + a_{k-1}^{(5)} \\ a_k^{(5)} = 8k + a_{k-1}^{(4)} \\ a_k^{(6)} = 4a_{k-1}^{(5)} + 2a_{k-2}^{(6)} \end{cases} \tag{3.1}$$

Step 1 builds the graph $G$ shown in Figure 3.1. We then, in step 2, initialize $d_1 = -\infty$, $d_2 = -\infty$, $d_3 = 2$, $d_4 = 0$, $d_5 = 1$, and $d_6 = -\infty$. We now enter step 3. Observe that

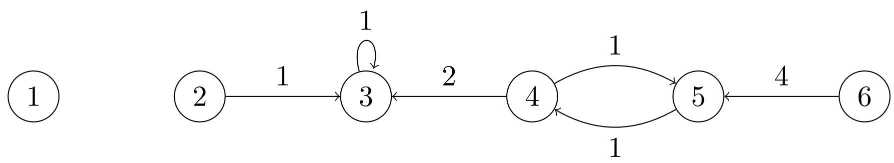

Figure 3.2: The graph $G'$ in our example

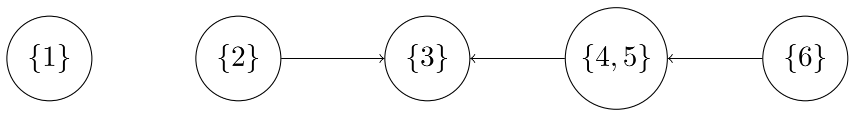

Figure 3.3: The graph $H$ in our example

vertex 6 in $G$ is in a circuit with an edge of weight 2. Also, observe that vertices 1 and 2 are each in more than one directed circuit. (One of the circuits for vertex 1 follows the loop on vertex 2; the other does not.) So, $W = \{1, 2, 6\}$. We set $d_1 = d_2 = d_6 = \infty$, and we obtain the graph $G'$ depicted in Figure 3.2.

We proceed now to steps 4 and 5. Each vertex of $G'$ is its own equivalence class except for 4 and 5, which are equivalent to each other. We obtain the graph $H$ shown in Figure 3.3. In step 6, we initialize the variables $d_{\{1\}} = \infty$, $d_{\{2\}} = \infty$, $d_{\{3\}} = 2$, $d_{\{4,5\}} = 1$, and $d_{\{6\}} = \infty$. We now move to step 7. One topological order of the vertices of $H$ is

$$\{6\}, \{4, 5\}, \{1\}, \{2\}, \{3\},$$

so we will process the vertices in the reverse of this order. When we process $\{3\}$, $d_{\{3\}}$ increases from 2 to 3, as vertex 3 in $G'$ has a loop. We also must increase $d_{\{4,5\}}$ to 3, as $3 = \max(1, 3)$. We then process $\{2\}$ and $\{1\}$, which causes no changes. When we reach $\{4, 5\}$, $d_{\{4,5\}}$ increases to 4, since 4 and 5 are in a cycle together in $G'$. The variables that started infinite cannot change, so the final values we obtain are $d_{\{1\}} = \infty$, $d_{\{2\}} = \infty$, $d_{\{3\}} = 3$, $d_{\{4,5\}} = 4$, and $d_{\{6\}} = \infty$. This tells us that any eventually positive solution to the positive recurrence system (3.1) will have $a_k^{(1)}$, $a_k^{(2)}$, and $a_k^{(6)}$ growing exponentially, $a_k^{(3)}$ growing cubically, and $a_k^{(4)}$ and $a_k^{(5)}$ growing quartically.

## 3.2 The Algorithm

In this section, we describe our algorithm for discovering positive-recurrent solutions to nested recurrences, which is implemented in our Maple package as the procedure `FindQgSolutions`. In Subsections 3.2.1, 3.2.2, and 3.2.3, we carefully describe and study how to discover period-$m$ positive-recurrent solutions to a basic recurrence. The process should resemble an attempt to prove the closed form to Golomb's solution (2.1) despite not yet knowing what form is desired. Then, Subsection 3.2.4 consists of a sample run through the algorithm. Finally, Subsection 3.2.5 discusses some possible relaxations on the inputs and outputs of the algorithm.

### 3.2.1 Input and Output

The primary input to the algorithm is a basic nested recurrence $L(n)$. It is perfectly reasonable to search for positive-recurrent solutions to other linear nested recurrences, or even nonlinear ones. But we fixate on basic recurrence here, as their simple form allows us to prove that our algorithm is correct. For discussions of what may happen with more general recurrences, see Subsection 3.2.5 and Chapter 11.

We also must guarantee that the algorithm terminates. This requires us to limit the search space in some way. A straightforward choice is to specify, as an input, the period $m$ of the positive-recurrent solutions we obtain. So, our input is a basic recurrence $L(n)$ and an integer $m \geq 1$ specifying the intended period.

The algorithm outputs a collection of positive-recurrent solutions that it discovers. Each item that it outputs actually specifies a family of solutions that all satisfy similar positive recurrence systems. More precisely, each item of the output consists of the following components:

1. A positive-recurrence system $\mathcal{P}$ containing some symbolic parameters

2. A set $\mathcal{C}$ of constraints on the symbolic parameters in $\mathcal{P}$

3. Values $\mathcal{V}$ for the parameters in $\mathcal{P}$ that satisfy all the constraints in $\mathcal{C}$

4. An initial condition `IC` that generates a positive-recurrent solution to $L(n)$ that satisfies $\mathcal{P}$ when the values $\mathcal{V}$ are substituted for the parameters.

The system $\mathcal{P}$, along with the constraints $\mathcal{C}$, defines a family of positive recurrence systems that might be realized by solutions to the recurrence $L(n)$. This should be thought of as a solution *template*. The values $\mathcal{V}$ then provide a certificate that the constraints are actually satisfiable. Finally, the initial condition `IC` certifies that the template is actually realizable in a solution.

We have already seen two examples of period-3 positive-recurrent solutions to the Hofstadter $Q$-recurrence, namely those of Golomb and Ruskey. Our algorithm, when given the $Q$-recurrence and $m = 3$ as input, outputs 12 items as described above. One of these outputs consists of the following items.

- Positive recurrence system $\mathcal{P}$, with symbolic parameters $\mu_1$ and $\mu_2$:

$$\begin{cases} a_k^{(0)} = a_{k-\frac{\mu_2}{3}}^{(0)} + a_{k-\frac{\mu_1}{3}}^{(0)} \\ a_k^{(1)} = \mu_1 \\ a_k^{(2)} = \mu_2 \end{cases}$$

- Constraints $\mathcal{C}$:
  - $\mu_1 > 0$
  - $\mu_2 > 0$
  - $\mu_1 \equiv 0 \pmod 3$
  - $\mu_2 \equiv 0 \pmod 3$
- Satisfying values $\mathcal{V}$: $\mu_1 = 3$, $\mu_2 = 3$
- Certifying initial condition: $\texttt{IC} = \langle 3, 3, 22, 3, 3, 22, 3, 3, 22, 3, 3 \rangle$.

This is an infinite family of solutions parametrized by pairs of positive multiples of three. The certificate given here is for $\mu_1 = 3$ and $\mu_2 = 3$. But, this family also contains Ruskey's solution, obtained from $\mu_1 = 3$ and $\mu_2 = 6$.

### 3.2.2 Steps of the Algorithm

We now describe the steps of the algorithm. Each step discussed in this subsection has a corresponding example application in Subsection 3.2.4.

#### Fixing the Behavior of the Subsequences

A positive-recurrent solution to $L(n)$ with period $m$ has the form

$$\begin{cases} L(mk) = a_k^{(0)} \\ L(mk+1) = a_k^{(1)} \\ L(mk+2) = a_k^{(2)} \\ \qquad\vdots \\ L(mk+(m-1)) = a_k^{(m-1)} \end{cases}$$

for some sequences $\left(a_k^{(0)}\right)$ through $\left(a_k^{(m-1)}\right)$. (For convenience, we index the interleaved sequences from zero in this context.) We define the following growth properties that these component sequences may have:

**Definition 3.5.**

- *Call* $\left(a_k^{(r)}\right)$ constant *if, for sufficiently large* $k$, $a_k^{(r)} = A$ *for some constant* $A$*.*
- *Call* $\left(a_k^{(r)}\right)$ linear *if, for sufficiently large* $k$, $a_k^{(r)} = Ak + B$ *for some constants* $A$ *and* $B$*.*
- *Call* $\left(a_k^{(r)}\right)$ superlinear *if* $a_k^{(r)} = \omega(k)$*.*
- *Call* $\left(a_k^{(r)}\right)$ standard linear *if* $a_k^{(r)} = mk + B$ *for some constant* $B$*.*
- *Call* $\left(a_k^{(r)}\right)$ steep linear *if* $a_k^{(r)} = Ak+B$ *for some constants* $A$ *and* $B$ *with* $A > m$*.*
- *Call* $\left(a_k^{(r)}\right)$ steep *if* $\left(a_k^{(r)}\right)$ *is either steep linear or superlinear.*

To start, we need to decide, for each of the $m$ component sequences, are we looking for a solution where that subsequence is *constant*, *standard linear*, or *steep*? To keep track of our choices, the algorithm stores variables $\lambda_0, \lambda_1, \ldots, \lambda_{m-1}$. We set $\lambda_r = 0$ if we decide that $\left(a_k^{(r)}\right)$ is to be constant, $\lambda_r = m$ if $\left(a_k^{(r)}\right)$ is standard linear, and $\lambda_r = \infty$ if $\left(a_k^{(r)}\right)$ is steep. In general, to perform an exhaustive search for positive-recurrent solutions, we iterate through the $3^m$ possible overall behaviors.

**Unpacking the Recurrence**

At this stage, we try to mirror the inductive step of the proof of Golomb's sequence's structure. That step consists of three parts; in general, this stage consists of $m$ similar parts. For each of the expressions $L(mk + r)$ with $0 \leq r < m$, we begin by expanding out the basic recurrence as

$$L(mk + r) = P(mk + r) + \sum_{i=1}^{d} \alpha_i L\left(mk + r - \beta_i - \sum_{j=1}^{d_i} L(mk + r - \gamma_{i,j})\right). \qquad (3.2)$$

This gives us a sum of expressions consisting of calls to $L$ themselves containing calls to $L$. Continuing to mirror the Golomb proof, we now try to use an inductive hypothesis to replace the inner calls to $L$. The $(i,j)^{th}$ inner call is of the form $L(mk'_{i,j} + r'_{i,j})$ with $0 \leq r'_{i,j} < m$, where either $k'_{i,j} < k$ or both $k'_{i,j} = k$ and $r'_{i,j} < r$. Specifically, we have $r'_{i,j} = (r - \gamma_{i,j}) \bmod m$ and $k'_{i,j} = k + \frac{r - \gamma_{i,j} - r'_{i,j}}{m}$.

Our inductive hypothesis comes from the behavior choices we make in the first stage of the algorithm. Observe that if a sequence $\left(a_k^{(r)}\right)$ is not steep, then $\left(a_k^{(r)}\right) = \lambda_r k + \mu_r$ for some constant $\mu_r$, where $\lambda_r$ is our variable tracking the growth of $\left(a_k^{(r)}\right)$. At this stage, introduce *symbols* $\mu_r$ for each of these constants. The inductive hypothesis is then:

If $k' < k$ or both $k' = k$ and $r' < r$, then:

- If $\left(a_k^{(r')}\right)$ is not steep, then $L(mk' + r') = \lambda_{r'} k + \mu_{r'}$.
- If $\left(a_k^{(r')}\right)$ is steep, then $L(mk' + r') \geq mk + r$.

For notational convenience, we also introduce a symbol $\mu_r$ if $\left(a_k^{(r)}\right)$ is steep and, in a sense, we "pretend" that $\left(a_k^{(r)}\right) = \lambda_r k + \mu_r$. This will allow us to treat the second part of the inductive hypothesis in-line.

Applying our inductive hypothesis to (3.2) yields

$$\begin{aligned} L(mk+r) &= P(mk+r) + \sum_{i=1}^{d} \alpha_i L\Bigg( mk + r - \beta_i \\ &\qquad - \sum_{j=1}^{d_i} \left( \lambda_{r'_{i,j}} \left( k + \frac{r - \gamma_{i,j} - r'_{i,j}}{m} \right) + \mu_{r'_{i,j}} \right) \Bigg) \\ &= P(mk+r) + \sum_{i=1}^{d} \alpha_i L\Bigg( \left( m - \sum_{j=1}^{d_i} \lambda_{r'_{i,j}} \right) k + r - \beta_i \\ &\qquad - \sum_{j=1}^{d_i} \mu_{r'_{i,j}} - \sum_{j=1}^{d_i} \lambda_{r'_{i,j}} \left( \frac{r - \gamma_{i,j} - r'_{i,j}}{m} \right) \Bigg) \end{aligned}$$

Now, for a given value of $i$, define

$$\ell_i = \frac{1}{m} \sum_{j=1}^{d_i} \lambda_{r'_{i,j}}.$$

The replacement yields the term

$$L\left( m\left(1 - \ell_i\right) k + r - \beta_i - \sum_{j=1}^{d_i} \mu_{r'_{i,j}} - \ell_i m \sum_{j=1}^{d_i} \frac{r - \gamma_{i,j} - r'_{i,j}}{m} \right).$$

If $\ell_i > 1$, then the coefficient on $k$ here is negative. This means that the index is negative for sufficiently large $k$. We assume inductively that such terms are zero. There are actually two sub-cases. If some sequence $\left(a_k^{(r'_{i,j})}\right)$ is steep, $\ell_i = \infty$ and the above term is of the form $L(-\infty k - c)$. In an inductive proof, we assume, at this point, that the terms of the steep sequence are larger than the indices they are being compared to. In this way, the infinite coefficient is convenient shorthand. If $1 < \ell_i < \infty$, then there are no steep sequences among the $\left(a_k^{(r'_{i,j})}\right)$, but $\ell_i$ of them are standard linear. Then, as long as $k$ is sufficiently large, the index is negative.

If $\ell_i = 1$, the unpacking gives

$$\begin{aligned} &L\left( r - \beta_i - \sum_{j=1}^{d_i} \mu_{r'_{i,j}} - m \sum_{j=1}^{d_i} \frac{r - \gamma_{i,j} - r'_{i,j}}{m} \right) \\ &= L\left( r - \beta_i - \sum_{j=1}^{d_i} \left( \gamma_{i,j} + r'_{i,j} - \mu_{r'_{i,j}} - r \right) \right). \end{aligned}$$

Such a term is a reference to a *particular* evaluation of $L$, so this term itself is a constant. Going forward, such terms are treated as additional symbolic parameters.

The only remaining case is where $\ell_i = 0$. Here, for the given value of $i$, each $\left(a_k^{(r'_{i,j})}\right)$ is constant, and each $\lambda_{r'_{i,j}} = 0$. So, we obtain an expression of the form $L(mk - c)$ for a symbolic constant $c$. Such expressions must themselves be handled inductively. We know that, in the $i^{th}$ term, the constant $c$ is $\beta_i - r + \sum_{j=1}^{d_i} \mu_{r'_{i,j}}$. Denote this quantity by $c_i$. We know that $L(mk - c_i) = L(mk''_i - r''_i)$ for $r''_i = (-c_i) \bmod m$ and $k''_i = k - \frac{c_i + r''_i}{m}$. The values of $r''_i$ and $k''_i$ depend on the congruence class of $c_i$ mod $m$. But, $c_i$ contains the symbols $\mu_{r'_{i,j}}$ for $1 \leq j \leq d_i$. So, we formally assign every $\mu_{r'_{i,j}}$ each of the $m$ possible congruences and iterate through all these possibilities. Once we select these congruence classes, if $\lambda_{r''_i} < \infty$, we replace $L(mk - c_i)$ by $\lambda_{r''_i} k'' + \mu_{r''_i}$. If $\lambda_{r''_i} = \infty$, we rewrite the expression as

$$L(mk - c_i) = L(mk'' + r''_i) = L\left(m\left(k - \frac{c_i + r''_i}{m}\right) + r''_i\right)$$

without eliminating the call to $L$.

The above process requires selecting congruences for the symbols $\mu_{r'_{i,j}}$ that appear in the constants $c_i$. These constants only appear when $\lambda_{r'_{i,j}} = 0$, so the only symbols for which we need to fix congruences are the symbols $\mu_r$ where $\left(a_k^{(r)}\right)$ is constant.

To summarize, we have now inductively "unpacked" the recurrence to the following form:

$$L(mk + r) = P(mk + r) + \sum_{\substack{1 \leq i \leq d \\ \ell_i = 1}} \alpha_i L\left(r - \beta_i - \sum_{j=1}^{d_i} \left(\gamma_{i,j} + r'_{i,j} - \mu_{r'_{i,j}} - r\right)\right) \tag{3.3}$$

$$+ \sum_{\substack{1 \leq i \leq d \\ \ell_i = 0 \\ \lambda_{r''_i} < \infty}} \alpha_i \left(\lambda_{r''_i}\left(k - \frac{c_i + r''_i}{m}\right) + \mu_{r''_i}\right) \tag{3.4}$$

$$+ \sum_{\substack{1 \leq i \leq d \\ \ell_i = 0 \\ \lambda_{r''_i} = \infty}} \alpha_i L\left(m\left(k - \frac{c_i + r''_i}{m}\right) + r''_i\right). \tag{3.5}$$

Line (3.4) can be further divided into a sum of $\alpha_i \mu_{r''_i}$ over those $i$ with $\lambda_{r''_i} = 0$ and of $\alpha_i \left(mk - c_i + r''_i + \mu_{r''_i}\right)$ over those $i$ with $\lambda_{r''_i} = m$.

We can now extract a system of nonhomogeneous linear recurrences from (3.3), (3.4), and (3.5). To accomplish this, we replace all expressions of the form $L(m\tilde{k}+\tilde{r})$ by $a_{\tilde{k}}^{(\tilde{r})}$. Specifically, for sequence $\left(a_k^{(r)}\right)$ we obtain

$$a_k^{(r)} = P(mk+r) + \sum_{\substack{1\le i\le d\\ \ell_i=1}} \alpha_i L\left(r - \beta_i - \sum_{j=1}^{d_i}\left(\gamma_{i,j} + r'_{i,j} - \mu_{r'_{i,j}} - r\right)\right) \tag{3.6}$$

$$+ \sum_{\substack{1\le i\le d\\ \ell_i=0\\ \lambda_{r''_i}<\infty}} \alpha_i \left(\lambda_{r''_i}\left(k - \frac{c_i + r''_i}{m}\right) + \mu_{r''_i}\right) \tag{3.7}$$

$$+ \sum_{\substack{1\le i\le d\\ \ell_i=0\\ \lambda_{r''_i}=\infty}} \alpha_i a^{(r''_i)}_{k-\frac{c_i+r''_i}{m}}. \tag{3.8}$$

The coefficients in the homogeneous part are $\alpha_i$ values. By assumption, these coefficients are all nonnegative. So, we have obtained a positive recurrence system, provided that the polynomials

$$\begin{aligned}\tilde{P}_r(k) := P(mk+r) + \sum_{\substack{1\le i\le d\\ \ell_i=1}} \alpha_i L\left(r - \beta_i - \sum_{j=1}^{d_i}\left(\gamma_{i,j} + r'_{i,j} - \mu_{r'_{i,j}} - r\right)\right)\\ + \sum_{\substack{1\le i\le d\\ \ell_i=0\\ \lambda_{r''_i}<\infty}} \alpha_i \left(\lambda_{r''_i}\left(k - \frac{c_i + r''_i}{m}\right) + \mu_{r''_i}\right)\end{aligned} \tag{3.9}$$

are all eventually nonnegative.

In a later step, we force the polynomials $\tilde{P}_r(k)$ to be eventually nonnegative, so we do actually obtain a positive recurrence system at this stage. For ease of notation going forward, denote this positive recurrence system by $\mathcal{P}$.

### Checking for Structural Consistency

The previous step has given us a positive recurrence system $\mathcal{P}$ that is eventually satisfied by the $\left(a_k^{(r)}\right)$'s. But, the structure of $\mathcal{P}$ may not be consistent with the original forms we assigned to the $\left(a_k^{(r)}\right)$ sequences. First, we check the following conditions:

- If $\left(a_k^{(r)}\right)$ is constant, the expression in $\mathcal{P}$ for $a_k^{(r)}$ should consist only of constants.

- If $\left(a_k^{(r)}\right)$ is standard linear, the expression in $\mathcal{P}$ for $a_k^{(r)}$ should be of the form $mk + c$ for some expression $c$ not containing any $k$.
- If $\left(a_k^{(r)}\right)$ is steep, the expression in $\mathcal{P}$ for $a_k^{(r)}$ should have at least one of the following:
  - A term $\sigma k$ with $\sigma > m$.
  - A reference to some steep $\left(a_k^{(r')}\right)$.

If any of these is violated for any $r$, there is no solution with the given $A$ values and congruence conditions, as the inductive step in an attempted proof of such a solution would fail when setting the inductively computed expressions equal to the target expressions.

If all of the above conditions are satisfied, we must determine the nature of each steep $\left(a_k^{(r)}\right)$. In particular, we need to determine if each one is steep linear or superlinear. As a bonus, we will be able to determine the degree of $\left(a_k^{(r)}\right)$ if it is a polynomial. Since $\mathcal{P}$ is a positive recurrence system, we can use the algorithm from Section 3.1 to accomplish precisely this task, provided we will start with an eventually positive initial condition. (We construct this initial condition later, on p. 41.) In running this algorithm, we may find that we actually do not have a solution, as the third case above includes expressions like $a_k^{(r)} = a_{k-1}^{(r)}$, which do not result in steep sequences. Again, if this happens, we can rule out a solution of the type currently under consideration with the congruences currently under consideration, as the inductive step of the proof would fail to verify the growth rates of the sequences.

If, during this step, we determine that there are no solutions of the current type, terminate this branch of the computation without returning anything.

**Building a Constraint Satisfaction Problem**

If our parameters produce a symbolic solution, we now know precisely what the structure of that particular solution must be. At this point, we need to see if such a solution can actually be realized.

In order to have a solution, we must check the following:

- If $\left(a_k^{(r)}\right)$ is constant, we must have $\mu_r > 0$. Otherwise, our solution would have infinitely many nonpositive values. This is not allowed, as we would then not be able to explicitly calculate terms of the sequence.

- If $\left(a_k^{(r)}\right)$ is constant, $\mu_r$ must equal our expression the right side of the expression for $a_k^{(r)}$ in $\mathcal{P}$.

- If $\left(a_k^{(r)}\right)$ is standard linear, $\mu_r$ must equal the constant term on the right side of the expression for $a_k^{(r)}$ in $\mathcal{P}$.

- If $\left(a_k^{(r)}\right)$ is steep linear, we may need a steepness constraint. This constraint is fairly complicated; we describe it below.

- Any constant that we have forced to have a certain congruence mod $m$ must actually have that congruence.

- Each polynomial $\tilde{P}_r(k)$ defined by Equation (3.9) must be eventually nonnegative in order to force $\mathcal{P}$ to be a positive recurrence system. This is automatically the case if the degree of $\tilde{P}_r$ is greater than zero. We know that $P$ is eventually nonnegative, and any new linear coefficient introduced in $\tilde{P}_r$ is one of the nonnegative $\alpha_i$ values. In the case that $\tilde{P}_r$ is constant, we need to require $\tilde{P}_r \geq 0$.

- For any two constants of the form $L(c_1)$ and $L(c_2)$ that appear, if $c_1 = c_2$ then $L(c_1)$ must equal $L(c_2)$.

- If constant $L(c)$ appears, then if $c \leq 0$ we must have $L(c) = 0$.

The last two of these restrictions gives a set of constraints of the form "If $X$ then $Y$" to check. For this reason, we call such constraints *conditional*. The rest of the constraints are unconditional constraints.

As mentioned above, constraining steep linear $\left(a_k^{(r)}\right)$'s to actually be steep requires a more complicated constraint. This stems from the fact that steep linear $\left(a_k^{(r)}\right)$'s can arise in three different ways. The steep linear $\left(a_k^{(r)}\right)$'s are a subset of the linear $\left(a_k^{(r)}\right)$'s. In terms of the algorithm in Section 3.1, the linear $\left(a_k^{(r)}\right)$'s are the ones for which $d_{\bar{r}} = 1$ when the algorithm terminates. The following are all the ways this could happen, in terms of the graphs $G'$ and $H$ in that algorithm.

1. The expression for $L(mk + r)$ is a degree-1 polynomial.

2. Vertex $r$ is not in a cycle in $G'$ and, when it came time to assign $d_{\bar{r}}$ its final value, the largest $d$ value in $H$ it pointed to was a 1.

3. Vertex $r$ is in a cycle in $G'$ and, when it came time to assign $d_{\bar{r}}$ its final value, the largest $d$ value in $H$ it pointed to was a 0 (or it pointed to no other vertices in $H$).

In Case 1, $\left(a_k^{(r)}\right)$ is steep linear if and only if the leading coefficient of that polynomial is greater than $m$. We already checked this in our structural consistency check, so if $\left(a_k^{(r)}\right)$ is linear because of Case 1, we need no steepness constraint. In Case 2, we have that $r$ is pointing to something else linear. But, our unpacking step would have removed all references to standard linear $\left(a_k^{(r')}\right)$'s. So, in Case 2, every $\left(a_k^{(r')}\right)$ that $\left(a_k^{(r)}\right)$ still refers to must be steep linear. This immediately forces $\left(a_k^{(r)}\right)$ itself to be steep linear without imposing any extra constraints.

This leaves only Case 3. In this case, $r$ is in a directed cycle in $G'$, say $r = r_0, r_1, r_2, \ldots, r_{t-1}, r_t = r$. Each of the corresponding sequences has an expression of the form

$$a_k^{(r_i)} = q_i + a_{k-e_i}^{(r_{i+1})}.$$

Repeated substitution yields the formula

$$a_k^{(r)} = \sum_{i=0}^{t-1} q_i + a_{k-\sum_{i=0}^{t-1} e_i}^{(r)}$$

for some constants $e_i$. We require that $\left(a_k^{(r)}\right)$ be steep. This will be accomplished if we have that

$$\sum_{i=0}^{t-1} q_i > m \sum_{i=0}^{t-1} e_i.$$

So, this is the steepness constraint we add in Case 3. In particular, we arrive at the same constraint for all sequences in a given equivalence class.

For notational convenience going forward, denote the system of constraints we obtain in this step by $\mathcal{C}$.

### Solving the Constraint Satisfaction Problem

As mentioned in Subsection 3.2.1, the systems $\mathcal{P}$ and $\mathcal{C}$ together represent a template for solutions to the recurrence $L(n)$. Given such a template, the inductive step of proving the structure of our sequence succeeds. Our goal is now to determine whether the template is actually satisfiable.

The system of unconditional constraints in $\mathcal{C}$ is almost an integer program. The following modifications can turn it into an integer program:

- Since all variables are integers, strict inequalities of the form $x > y$ can be made loose by replacing them by the equivalent inequalities $x \geq y + 1$.
- Congruence constraints can be converted to equality constraints via the introduction of auxiliary variables. Namely, $x \equiv y \pmod{m}$ is the same constraint as $x = Km + y$, where $K$ is a new auxiliary variable.

Furthermore, the conditional constraints can be incorporated into the integer program. For each constraint of the form "If $(c_1 = c_2)$, then $(L(c_1) = L(c_2))$," we consider three cases. (If one fails, we try the next one.)

- Add the constraints $c_1 = c_2$ and $L(c_1) = L(c_2)$.
- Add the constraint $c_1 \leq c_2 - 1$.
- Add the constraint $c_1 \geq c_2 + 1$.

And, for each constraint of the form "If $(c \leq 0)$, then $(L(c) = 0)$," we consider two cases.

- Add the constraints $c \leq 0$ and $L(c) = 0$.
- Add the constraint $c \geq 1$.

Maple has a built-in procedure, `LPSolve`, that can satisfy *linear* integer programs. Since we are only considering linear nested recurrences, the integer program we obtain is, in fact, linear. Integer linear programming is an NP-hard problem [24]. But, experimentally, the instances that arise in this context seem to be not very hard. Heuristically guessing $\mu_r$ values that are far apart seems to make satisfying the constraints be a quick process. In particular, the constraints we obtain are typically satisfiable unless there is some obvious reason why they should not be. If it turns out the constraint system is not satisfiable, there is no solution in this branch. In that case, terminate the branch without returning anything.

Going forward, denote the value assignment we obtain here by $\mathcal{V}$. Now that we have specific values for the parameters in our template, we can carry out the inductive step of a Golomb-like proof using specific values. All that remains is the base case.

**Constructing an Initial Condition**

We now have an eventual solution to our recurrence given by the template $\mathcal{P}$ and $\mathcal{C}$ along with the value assignments $\mathcal{V}$. These together suffice to carry out the inductive step of a Golomb-like proof of the structure of a solution to the recurrence $L(n)$. But, any proof by induction needs a base case. Much like we do in the proof of Golomb's solution (p. 10), we must answer the question: "What do we assume in the inductive step?" There are a few different types of assumptions that may be made:

1. There are sometimes constraints involving terms of the form $L(c)$ for some constant $c$.

2. The rewriting of $L(mk+r)$ refers to some terms of the form $L(mk'+r')$ for various $k' \leq k$ and various $r'$. Such rewriting is only valid if the referenced terms actually have the values they are assumed to have. Let $k_0$ be the minimum such $k'$ that was referenced for any $r$. The inductive step is only guaranteed to be valid if $k_0 > \left\lfloor \frac{c_0}{m} \right\rfloor$, as there may be anomalous terms at or below $L(c_0)$.

3. Also, we sometimes substitute $\infty$ into expressions when rewriting $L(mk+r)$. Each time we do this, we are assuming that the terms in a steep sequence $\left(a_k^{(r')}\right)$ are greater than the indices referencing these terms. This is eventually true of any steep sequence, but, it may fail for finitely many initial values.

4. Similarly, rewriting $L(mk+r)$ sometimes includes expressions with negative coefficients on $m$. The inductive step is only valid if $k$ is large enough that all such

referenced indices are nonpositive.

We now use the above observations to construct an initial condition that allows the induction to be carried out safely. Let $c_0$ denote the maximum positive constant $c$ so that $L(c)$ appears in $\mathcal{V}$, or let $c_0 = 0$ if no such positive $c$ exists. Let $\gamma$ denote the maximum of the constants $\gamma_{i,j}$ appearing in the recurrence $L(n)$, and, similarly, let $\beta$ denote the maximum of the $\beta_i$. Then, let $\mu$ denote the maximum value of $\mu_r$ in $\mathcal{V}$ for any $r$ with $\left(a_k^{(r)}\right)$ constant. Also, define

$$\kappa = \left\lceil \max\left\{\frac{-\epsilon m}{c} : L(\epsilon k - c) \text{ appears in the unpacking}\right\}\right\rceil.$$

Then, let $\nu$ be the next integer greater than or equal to $c_0 + 2m + \max(\gamma, \beta + \mu, \kappa)$ that is congruent to $-1 \bmod m$. Define an initial condition `IC` as follows:

1. Start with $\texttt{IC} = \langle L(1), L(2), \dots, L(\nu)\rangle$, a generic initial condition of length $\nu$.

2. Any values of $L(c)$ defined by $\mathcal{V}$ should be placed in those positions in `IC`. (For example if $\mathcal{V}$ contains an assignment of $L(2) = 7$, then the second entry in `IC` should become 7.)

3. If $\left(a_k^{(r)}\right)$ is not steep, then for each index $1 \le mk+r \le \nu$, set $L(mk+r) = \lambda_r k + \mu_r$ in `IC`, unless this term was already set by item 2.

4. Any remaining symbolic terms in `IC` are in steep positions. Set each of these terms to the value $2\nu$.

We have the following claim:

**Claim 3.6.** *The initial condition `IC` obtained above generates a positive-recurrent solution to the recurrence $L(n)$ with eventual behavior described by $\mathcal{P}$ and $\mathcal{V}$.*

*Proof.* We already know that the inductive step in proving this eventual behavior succeeds. So, it suffices to show that `IC` provides a valid base case for that induction. That is, we must show that the inductive step works for all indices greater than $\nu$, even if a term in the induction references a term in `IC`.

Let $mk+r > \nu$. Importantly, this means that $k \geq \frac{\nu+1}{m}$. First, suppose we encounter an inner expression $L(mk'+r')$ when unpacking $L(mk+r)$ with $mk'+r' \leq \nu$. We know that $mk'+r' \geq mk+r-\gamma$. Using the fact that $r-r' \geq 1-m$, we can calculate

$$\begin{aligned}
k' &\geq k + \frac{r-\gamma-r'}{m} \\
&\geq \frac{\nu+1}{m} + \frac{r-\gamma-r'}{m} \\
&\geq \frac{c_0+2m+1+r-r'}{m} \\
&\geq \frac{c_0+m+2}{m} \\
&> \frac{c_0}{m} + 1 \\
&> \left\lfloor \frac{c_0}{m} \right\rfloor .
\end{aligned}$$

So, rewriting $L(mk'+r')$ does not refer to any index with a forced value. This leaves two possibilities:

$\left(a_k^{(r')}\right)$ **is not steep:** In this case, item 3 assures us that the value of $L(mk'+r')$ is $\lambda_{r'}k' + \mu_{r'}$. This is the value it is assumed to be in the inductive hypothesis.

$\left(a_k^{(r')}\right)$ **is steep:** In this case, item 4 assures us that the value of $L(mk'+r')$ is $2\nu$. Since $\nu \geq mk'+r' \geq mk+r-\gamma$, we have that $mk+r \leq \nu+\gamma \leq 2\nu$. So, $L(mk'+r') \geq mk+r$, which is the same assumption we make in the inductive hypothesis.

Now, suppose we encounter an inner expression $L(mk'+r')$ when unpacking $L(mk+r)$ that we replace by $\mu_{r'}$ because $\left(a_k^{(r')}\right)$ is constant. The outer expression we are left with is then $L(mk+r-\beta_i-\mu_{r'})$ for some $i$. Suppose we have this $mk+r-\beta_i-\mu_{r'} \leq \nu$. We know that $mk+r-\beta_i-\mu_{r'} = mk''+r''$ for some $0 \leq r'' < m$. Solving for $k''$ gives

$k'' = k + \frac{r-\beta_i-\mu_{r'}-r''}{m}$. Using the fact that $r - r'' \geq 1 - m$, we can calculate

$$\begin{aligned} k'' &\geq k + \frac{r - \beta_i - \mu_{r'} - r''}{m} \\ &\geq \frac{\nu + 1}{m} + \frac{r - \beta - \mu - r''}{m} \\ &\geq \frac{c_0 + 2m + 1 + r - r''}{m} \\ &\geq \frac{c_0 + m + 2}{m} \\ &> \left\lfloor \frac{c_0}{m} \right\rfloor . \end{aligned}$$

So, rewriting $L(mk'' + r'')$ does not refer to any index with a forced value. This leaves two possibilities:

$\left(a_k^{(r'')}\right)$ **is not steep:** In this case, item 3 assures us that the value of $L(mk'' + r'')$ is $\lambda_{r''}k' + \mu_{r''}$. This is the value it is assumed to be in the inductive hypothesis.

$\left(a_k^{(r'')}\right)$ **is steep:** In this case, item 4 assures us that the value of $L(mk'' + r'')$ is $2\nu$. Since $\nu \geq mk'' + r'' \geq mk + r - \beta - \mu$, we have that $mk + r \leq \nu + \beta + \mu \leq 2\nu$. So, $L(mk'' + r'') \geq mk + r$. At this stage in the unpacking, we did not evaluate the outer expression. But, this calculation ensures us that our initial condition values are large enough for all remaining terms of $\left(a_k^{(r'')}\right)$ to be large enough.

Finally, suppose we encounter an inner expression $L(\epsilon k - c)$ with $-\infty < \epsilon < 0$ when unpacking $L(mk + r)$. Since $\nu \geq \kappa \geq \frac{\epsilon m}{c}$, we have

$$k \geq \frac{\nu + 1}{m} \geq \frac{\frac{\epsilon m}{c}}{m} = \frac{\epsilon}{c}.$$

So, $\epsilon k - c \leq 0$, which means that, in the induction, our assumption that $L(\epsilon k - c) = 0$ is correct. $\square$

The initial condition that we construct here depends on the values $\mathcal{V}$. But, there are probably infinitely many possible choices of $\mathcal{V}$ that could have satisfied $\mathcal{C}$. This method of constructing an initial condition works for any such choice of $\mathcal{V}$.

Note that the initial condition we obtain in this step is not "optimal." There may be shorter initial conditions that generate the same eventual sequence, and some of the constraints on symbolic terms in the initial condition may be able to be relaxed somewhat. Our Maple implementation is a more complicated variation on this construction; it attempts to produce optimal initial conditions.

At this point, this branch of the algorithm is complete. Add the 4-tuple $(\mathcal{P}, \mathcal{C}, \mathcal{V}, \mathtt{IC})$ to the collection of objects to be returned and terminate this branch of the algorithm.

### 3.2.3 Correctness, Termination, and Complexity

We claim that the algorithm described throughout Subsection 3.2.2 is correct and terminates. By "correct," we mean that every item it outputs actually does specify a positive-recurrent solution. The proof of correctness is scattered throughout the descriptions of the steps themselves. The start, on p. 32, just requires iterating through $3^m$ cases. In the next step (p. 33), we obtain a parametrized positive recurrence system $\mathcal{P}$ that our solution will satisfy. This step involves iterating through congruence choices for each $\mu_r$ with $\left(a_k^{(r)}\right)$ constant. So, we only have to spawn finitely many branches in this step (at most $m^m$ of them). Then (p. 36), we use the graph algorithm from Section 3.1 as a subroutine. That algorithm's correctness is proved by Claim 3.4, and it clearly terminates. In the next step (p. 37), we determine what constraints $\mathcal{C}$ we require. Each of the constraints derived there is derived by a finite process. Then, we satisfy the constraints (p. 40). Here, we use established integer programming algorithms as subroutines, as well as some backtracking through the conditional constraints. Finally (p. 41) we, in polynomial time, provably find an initial condition of bounded length.

We can actually make a stronger claim about the correctness of our algorithm. Suppose the recurrence $L(n)$ has a positive-recurrent solution satisfying positive recurrence system $\mathcal{P}$. As long as every component sequence of $\mathcal{P}$ is constant, standard linear, or steep, our algorithm will have discovered a family of solutions including this solution. At some point, the algorithm will have guessed the correct behaviors for the subsequences and the correct congruences for the constants. Since the recurrence, by assumption, has a solution with those parameters, the resulting constraint system will have definitely been satisfiable. The satisfying values then lead to an initial condition.

In short, the only possible positive-recurrent solutions that our algorithm can miss are ones with a component sequence that is not constant, not standard linear, and not steep. By the graph algorithm in Section 3.1, each component sequence either grows polynomially or exponentially. So, the only types of solutions that could have been missed are:

- Solutions with a linear component sequence whose slope lies strictly between 0 and 1
- Solutions with a quasilinear component sequence that is not itself constant or linear.

Our algorithm can still find solutions of the second type, but it would find them for a larger input $m$. No solutions of the first type are known for any basic recurrence with $P(n) = 0$ and each $d_i = 1$. For an example of such a solution to a basic recurrence with $d = 2$, $d_1 = 2$, and $d_2 = 3$, see 3.3.2 at the end of this chapter.

We now analyze the running time of our algorithm. The first step (p. 32), contributes a factor of $3^m$ to the worst-case complexity. Then, the next step (p. 33) contributes a worst-case factor of $m^m$. Most of the computation of the following step (p. 36) involves running the graph algorithm from Section 3.1. This is a polynomial time subroutine, as it is based on topological sort and graph search algorithms. The only remaining part that may not run in polynomial time is the constraint satisfaction on p. 40. In the worst case, this could require checking $3^c$ integer programs for feasibility, where $c$ is the number of conditional constraints. The complexity of each such check is the complexity of integer programming, which, as we mentioned, is an NP-hard problem. We can crudely bound the total number of constraints by $4m + md + \binom{md}{2}$ as follows:

- Each $0 \leq r < m$ contributes at most four constraints:
    - At most one positivity constraint on $\mu_r$
    - At most one equality constraint for $\mu_r$ or steepness constraint
    - At most one congruence constraint on $\mu_r$
    - At most one non-negativity constraint on $\tilde{P}_r(k)$

- There are at most $md$ additional symbolic constants of the form $L(c)$ for some $c$.
- Each such constant contributes at most one constraint of the form "If $c \leq 0$, then $L(c) = 0$."
- Each such pair of constants contributes at most one constraint of the form "If $c_1 = c_2$, then $L(c_1) = L(c_2)$."

So, the total number of constraints is at most $4m + md + \binom{md}{2}$. This implies that the worst-case complexity of the algorithm is at most

$$O\left((3m)^m \cdot 3^{4m+md+\binom{md}{2}} f\left(4m + md + \binom{md}{2}\right)\right),$$

where $f(n)$ is the complexity of our integer programming instances on at most $n$ variables and $n$ constraints.

This is a terrible theoretical running time, but, in practice, the algorithm is usable. First, if the user is only interested in certain sequence behaviors or certain congruences for constants, the first exponential terms can be reduced. Second, many cases that do not result in a solution will fail before attempting to satisfy any integer programs. Third, the integer programming step can stop once we find a solution; we do not need to continue branching all the way through the conditional constraints. Also, we are only interested in *finding one* feasible point in the domain of the integer program; general integer programming involves optimizing some objective function over *all* feasible points. This fact, along with the specific form of our integer programs, makes solving these programs not too hard in practice. Additionally, the bound we give on the number of constraints is very crude, and the constraints can oftentimes be reduced using some logical rules before even beginning to branch. The Maple implementation makes use of these and a few other optimizations.

### 3.2.4 An Example

We follow the steps of the algorithm by applying it to the following input:

**Recurrence:** $B_R(n) = B_R(n - B_R(n-1)) + B_R(n - B_R(n-2)) + B_R(n - B_R(n-3))$.

**Period:** $m = 4$

We choose this particular example because it will illustrate many facets of the algorithm without becoming too unwieldy.

### Fixing the Behavior of the Subsequences

We seek a solution of the form

$$\begin{cases} B_R(4k) = a_k^{(0)} \\ B_R(4k+1) = a_k^{(1)} \\ B_R(4k+2) = a_k^{(2)} \\ B_R(4k+3) = a_k^{(3)}. \end{cases}$$

The full algorithm would now iterate through all $3^4$ possibilities for $\lambda_0, \lambda_1, \lambda_2, \lambda_3 \in \{0, 4, \infty\}$. Going forward, we treat only the case $\lambda_0 = \infty$, $\lambda_1 = 4$, $\lambda_2 = 0$, and $\lambda_3 = 0$. That is, we are seeking a solution with $\left(a_k^{(0)}\right)$ steep, $\left(a_k^{(1)}\right)$ standard linear, and the other two sequences constant.

### Unpacking the Recurrence

Since $\left(a_k^{(2)}\right)$ and $\left(a_k^{(3)}\right)$ are constant, we need to decide the congruence classes of $\mu_2$ and $\mu_3$ mod 4. There are $4^2 = 16$ possibilities here to check. We treat only the case where $\mu_2 \equiv 0 \pmod{4}$ and $\mu_3 \equiv 3 \pmod{4}$; the full algorithm would treat all 16 cases. Under these assumptions, here is how the recurrence unpacks:

$$\begin{aligned} B_R(4k) &= B_R(4k - B_R(4k-1)) + B_R(4k - B_R(4k-2)) + B_R(4k - B_R(4k-3)) \\ &= R\left(4k - a_{k-1}^{(3)}\right) + R\left(4k - a_{k-1}^{(2)}\right) + R\left(4k - a_{k-1}^{(1)}\right) \\ &= B_R(4k - 0\,(k-1) - \mu_3) + B_R(4k - 0\,(k-1) - \mu_2) \\ &\quad + B_R(4k - 4\,(k-1) - \mu_1) \\ &= B_R(4k - \mu_3) + B_R(4k - \mu_2) + B_R(4 - \mu_1) \\ &= a^{(1)}_{k - \frac{\mu_3+1}{4}} + a^{(0)}_{k - \frac{\mu_2}{4}} + B_R(4 - \mu_1) \\ &= 4\left(k - \frac{\mu_3+1}{4}\right) + \mu_1 + a^{(0)}_{k - \frac{\mu_2}{4}} + B_R(4 - \mu_1) \end{aligned}$$

$$= 4k - \mu_3 - 1 + \mu_1 + a^{(0)}_{k-\frac{\mu_2}{4}} + B_R(4-\mu_1).$$

$$\begin{aligned}
B_R(4k+1) &= B_R(4k+1-B_R(4k)) + B_R(4k+1-B_R(4k-1)) \\
&\quad + B_R(4k+1-B_R(4k-2)) \\
&= R\Big(4k+1-a^{(0)}_k\Big) + R\Big(4k+1-a^{(3)}_{k-1}\Big) + R\Big(4k+1-a^{(2)}_{k-1}\Big) \\
&= B_R(4k+1-\infty k-\mu_0) + B_R(4k+1-0\,(k-1)-\mu_3) \\
&\quad + B_R(4k+1-0\,(k-1)-\mu_2) \\
&= B_R(-\infty k+1-\mu_0) + B_R(4k+1-\mu_3) + B_R(4k+1-\mu_2) \\
&= 0 + a^{(2)}_{k-\frac{\mu_3+1}{4}} + a^{(1)}_{k-\frac{\mu_2}{4}} \\
&= 0\left(k-\frac{\mu_3+1}{4}\right) + \mu_2 + 4\left(k-\frac{\mu_2}{4}\right) + \mu_1 \\
&= \mu_2 + 4k - \mu_2 + \mu_1 \\
&= 4k + \mu_1.
\end{aligned}$$

$$\begin{aligned}
B_R(4k+2) &= B_R(4k+2-B_R(4k+1)) + B_R(4k+2-B_R(4k)) \\
&\quad + B_R(4k+2-B_R(4k-1)) \\
&= R\Big(4k+2-a^{(1)}_k\Big) + R\Big(4k+2-a^{(0)}_k\Big) + R\Big(4k+2-a^{(3)}_{k-1}\Big) \\
&= B_R(4k+2-4k-\mu_1) + B_R(4k+2-\infty k-\mu_0) \\
&\quad + B_R(4k+2-0\,(k-1)-\mu_3) \\
&= B_R(2-\mu_1) + B_R(-\infty k+2-\mu_0) + B_R(4k+2-\mu_3) \\
&= B_R(2-\mu_1) + 0 + a^{(3)}_{k-\frac{\mu_3+1}{4}} \\
&= B_R(2-\mu_1) + 0\left(k-\frac{\mu_3+1}{4}\right) + \mu_3 \\
&= B_R(2-\mu_1) + \mu_3.
\end{aligned}$$

$$\begin{aligned}
B_R(4k+3) &= B_R(4k+3-B_R(4k+2)) + B_R(4k+3-B_R(4k+1)) \\
&\quad + B_R(4k+3-B_R(4k)) \\
&= R\Big(4k+3-a^{(2)}_k\Big) + R\Big(4k+3-a^{(1)}_k\Big) + R\Big(4k+3-a^{(0)}_k\Big)
\end{aligned}$$


Figure 3.4: The graph $G = G'$ at this stage of the algorithm's execution

$$
\begin{aligned}
&= B_R(4k+3-0k-\mu_2) + B_R(4k+3-4k-\mu_1) \\
&\quad + B_R(4k+3-\infty k-\mu_0) \\
&= B_R(4k+3-\mu_2) + B_R(3-\mu_1) + B_R(-\infty k+3-\mu_0) \\
&= a^{(3)}_{k-\frac{\mu_2}{4}} + B_R(3-\mu_1) + 0 \\
&= 0\left(k - \frac{\mu_2}{4}\right) + \mu_3 + B_R(3-\mu_1) \\
&= \mu_3 + B_R(3-\mu_1).
\end{aligned}
$$

From this unpacking, we obtain the positive recurrence system

$$
\mathcal{P} = \begin{cases}
a^{(0)}_k = 4k - \mu_3 - 1 + \mu_1 + a^{(0)}_{k-\frac{\mu_2}{4}} + B_R(4-\mu_1) \\
a^{(1)}_k = 4k + \mu_1 \\
a^{(2)}_k = B_R(2-\mu_1) + \mu_3 \\
a^{(3)}_k = \mu_3 + B_R(3-\mu_1).
\end{cases}
$$

**Checking for Structural Consistency**

We now verify successfully that $\left(a^{(1)}_k\right)$ is standard linear (the expression we obtained for $B_R(4k+1)$ is $4k+\mu_1$) and that $\left(a^{(2)}_k\right)$ and $\left(a^{(3)}_k\right)$ are constant (expressions $B_R(2-\mu_1)+\mu_3$ and $\mu_3 + B_R(3-\mu_1)$ respectively). We then run our graph algorithm on $\mathcal{P}$. The graph $G$, depicted in Figure 3.4, consists of four vertices. Vertex 0 has a loop with weight 1; the other three vertices are isolated. We initialize $d_0 = 1$, $d_1 = 1$, $d_2 = 0$, and $d_3 = 0$. Step 3 of that algorithm does not affect any of the vertices, so $G' = G$. Similarly, $\sim$ has no nontrivial relations, so $H \cong G$ via the isomorphism $i \leftrightarrow \{i\}$. When we process vertex $\{0\}$ in $H$, we set $d_{\{0\}} = 2$, and this is the only change made in Step 7. So, we obtain that $\left(a^{(0)}_k\right)$ is quadratic.

**Building a Constraint Satisfaction Problem**

In our example, we obtain the following constraint system $\mathcal{C}$:

- $\left(a_k^{(0)}\right)$ is superlinear, so there are no constraints associated to it.
- $\left(a_k^{(1)}\right)$ is standard linear, and we have $B_R(4k+1) = 4k + \mu_1$. This gives us the constraint $\mu_1 = \mu_1$. (This constraint is tautological, but this is okay.)
- $\left(a_k^{(2)}\right)$ is constant, and we have $B_R(4k+2) = B_R(2-\mu_1) + \mu_3$. This gives us the following constraints:
  - $\mu_2 > 0$
  - $\mu_2 = B_R(2-\mu_1) + \mu_3$.
- $\left(a_k^{(3)}\right)$ is constant, and we have $B_R(4k+3) = \mu_3 + B_R(3-\mu_1)$. This gives us the following constraints:
  - $\mu_3 > 0$
  - $\mu_3 = \mu_3 + B_R(3-\mu_1)$.
- Our congruence constraints are
  - $\mu_2 \equiv 0 \pmod 4$
  - $\mu_3 \equiv 3 \pmod 4$.
- The polynomials $\tilde{P}_2(k) = B_R(2-\mu_1) + \mu_3$ and $\tilde{P}_3(k) = \mu_3 + B_R(3-\mu_1)$ are constant. This gives us the following constraints:
  - $B_R(2-\mu_1) + \mu_3 \geq 0$
  - $\mu_3 + B_R(3-\mu_1) \geq 0$.

  These constraints end up being redundant, but, like the tautological constraint above, this is fine.
- Our conditional constraints are
  - If $2-\mu_1 = 3-\mu_1$, then $B_R(2-\mu_1)$ must equal $B_R(3-\mu_1)$.

  - If $2-\mu_1 \leq 0$, then $B_R(2-\mu_1) = 0$.
  - If $3-\mu_1 \leq 0$, then $B_R(3-\mu_1) = 0$.

One of these constraints is vacuously true, but, again, this is allowed.

### Solving the Constraint Satisfaction Problem

At this point, we call an integer program solver to find assignments that satisfy the constraints $\mathcal{C}$ or to determine that no such assignments exist. In our example, the following assignments $\mathcal{V}$ satisfy the constraints $\mathcal{C}$:

- $\mu_1 = 0$
- $\mu_2 = 4$
- $\mu_3 = 3$
- $B_R(2) = 1\,(= B_R(2-\mu_1))$
- $B_R(3) = 0\,(= B_R(3-\mu_1))$.

This means we have the following eventual solution:

$$\begin{cases} B_R(4k) = 4k-3-1+0+a_{k-1}^{(0)} + B_R(4) = B_R(4k-4)+4k+B_R(4)-4 \\ B_R(4k+1) = 4k \\ B_R(4k+2) = 4 \\ B_R(4k+3) = 3. \end{cases}$$

These constraints have other satisfying values, and each other satisfaction leads to another eventual solution to the recurrence.

### Constructing an Initial Condition

All that remains is to construct an initial condition `IC` that generates the eventual solution we have found. We first need to determine the value of $\nu$ that we described on p. 41. The assignments $\mathcal{V}$ give values for $B_R(2)$ and $B_R(3)$, so $c_0 = 3$. Also, we have $\mu_2 = 4$ and $\mu_3 = 3$, so $\mu = 4$. The recurrence tells us that $\beta = 0$ and $\gamma = 3$, and the

| | Index | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| Start | $B(1)$ | $B(2)$ | $B(3)$ | $B(4)$ | $B(5)$ | $B(6)$ | $B(7)$ | $B(8)$ | $B(9)$ | $B(10)$ | $B(11)$ | $B(12)$ | $B(13)$ | $B(14)$ | $B(15)$ |
| $\mathcal{V}$ | $B(1)$ | 1 | 0 | $B(4)$ | $B(5)$ | $B(6)$ | $B(7)$ | $B(8)$ | $B(9)$ | $B(10)$ | $B(11)$ | $B(12)$ | $B(13)$ | $B(14)$ | $B(15)$ |
| $\lambda_r k + \mu_r$ | 0 | 1 | 0 | $B(4)$ | 4 | 4 | 3 | $B(8)$ | 8 | 4 | 3 | $B(12)$ | 12 | 4 | 3 |
| Steep | 0 | 1 | 0 | 30 | 4 | 4 | 3 | 30 | 8 | 4 | 3 | 30 | 12 | 4 | 3 |

Table 3.1: Creation of initial condition `IC`

other input is $m = 4$. Also, we do not encounter any expressions with finite negative coefficients on $m$, so $\kappa = 0$. So, $\nu \geq c_0 + 2m + \max(\gamma, \beta + \mu, \kappa) = 3 + 8 + 4 = 15$. Since $15 \equiv -1 \pmod 4$, $\nu$ actually equals 15.

Table 3.1 describes how `IC` is built. We begin with a generic initial condition of length $\nu = 15$. The assignment $\mathcal{V}$ forces $B_R(2) = 1$ and $B_R(3) = 0$, so we assign those values of the initial condition first. Then, we fill the remaining non-steep indices (those indices that are not divisible by four) with values according to the eventual formulas for those subsequences. Finally, we substitute $30 = 2\nu$ for each remaining symbol. This gives us the initial condition

$$\mathtt{IC} = \langle 0, 1, 0, 30, 4, 4, 3, 30, 8, 4, 3, 30, 12, 4, 3 \rangle .$$

The resulting sequence is essentially A268368 in OEIS; that sequence results from optimizing the initial condition [31].

#### Output of Example

We have finished exploring one branch of the algorithm's execution on input $B_R(n)$ and $m = 4$. This branch successfully found a family of positive-recurrent solutions to $B_R(n)$ with period 4, and it adds $(\mathcal{P}, \mathcal{C}, \mathcal{V}, \mathtt{IC})$ to the collection of items to output. After traversing all of the branches, the algorithm outputs a total of 36 items of this form, corresponding to 36 different infinite families of period-4 positive-recurrent solutions to the recurrence $B_R(n)$.

### 3.2.5 Possible Generalizations

It makes perfect sense to search for positive-recurrent solutions to general linear nested recurrences, rather than just to basic ones. The algorithm we have described applies to general linear recurrences with almost no modification. We restrict our analysis to basic recurrences as we are able to give correctness and termination guarantees in this context. The primary issues that arise when working with non-basic linear recurrences are the following:

- In the analysis of the second step of the algorithm (p. 33), we discover that the $\mu_r$'s for which we need to specify congruences are precisely the ones for which $\left(a_k^{(r)}\right)$ is constant. This is no longer true for general linear recurrences; congruences for $\mu_r$ can matter here even when $\left(a_k^{(r)}\right)$ is standard linear.

- With a basic recurrence, it is easy to assert that the recurrence system we obtain at the end of the second step is a positive recurrence system. This is much harder to guarantee in general with a non-basic recurrence, but specific examples often do produce positive recurrence systems.

- The procedure (p. 41) for finding initial conditions involves a constant $\nu$ that depends on constants $\mu$, $\beta$, $\gamma$, $\kappa$, and $c_0$. The first three of these are defined assuming we have a basic recurrence. A similar procedure can work for specific nonbasic recurrences, but it is more difficult to describe a generic procedure.

The procedure `FindQgSolutions` in the Maple package, keeping these difficulties in mind, accepts any linear recurrence with one level of nesting as an input. The output is only guaranteed to be correct in the basic case, but the implementation generally succeeds in the nonbasic case also.

As mentioned in 3.2.3, we do not concern ourselves with component sequences intermediate in growth between constant and standard linear. These seem to be uncommon and can be harder to analyze. But, the Maple package can sometimes handle these, as long as the user explicitly asks it to.

## 3.3 Findings

### 3.3.1 Hofstadter $Q$-Recurrence

As our example (3.2.4) illustrates, it is possible to obtain quasi-quadratic solutions to Hofstadter-like recurrences. This begs the question as to whether higher degree quasipolynomials are possible. If they are, the algorithm can find them. It turns out that there are quasipolynomial solutions to the Hofstadter $Q$-recurrence of every positive degree [12]. (See Chapter 4 for details of this construction.) In addition, there are solutions that include both quadratic and exponential subsequences, such as the sequence obtained from the Hofstadter $Q$-recurrence with $\langle 9, 0, 0, 0, 7, 9, 9, 10, 4, 9, 9, 3\rangle$ as the initial condition [31, A275153]. It is likely that a construction similar to the one for arbitrary degree quasipolynomials [12] will also lead to examples including higher degree polynomials along with exponentials. There are also solutions to the Hofstadter $Q$-recurrence with linear subsequences with slopes greater than 1, and such subsequences can be obtained by any of the three ways mentioned on p. 39.

- The length-45 initial condition

  $$\langle 0, 4, -40, -9, 8, -8, 7, 1, 5, 13, -24, -1, 8, 8, 8, 1, 5, 13, -8, 7, 8, 8, 23, 1, 5, 13, 8, 15, 8, 16, 31, 1, 5, 13, 24, 23, 8, 24, 39, 1, 5, 13, 40, 31, 8\rangle$$

  leads to a period-8 solution with $Q(8n + 3) = 16n - 40$. This is the case because unpacking $Q(8n + 3)$ involves adding two standard linear terms together [31, A275361].

- The length-16 initial condition

  $$\langle -9, 2, 9, 2, 0, 7, 9, 10, 3, 0, 2, 9, 2, 9, 9, 9\rangle$$

  leads to a period-9 solution where $Q(9n + 2)$ and $Q(9n + 8)$ both have slope 10. But, $Q(9n + 2)$ has slope 10 because unpacking it yields $Q(9n - 1)$ plus a constant [31, A275362].

- In the previous example, $Q(9n + 8)$ has slope 10 because unpacking it results in

$10 + Q(9n - 1)$. This appears to be, by far, the most common way steep linear solutions arise.

We have used our algorithm to fully explore positive-recurrent solution families to the Hofstadter $Q$-recurrence with small periods. Given a solution to the Hofstadter $Q$-recurrence, any shift of it is also a solution, since the recurrence only depends on the relative indices of the terms (and not the absolute indices). So, solution families found by the algorithm can be considered as-is or modulo the shifting operation. Our algorithm finds no period-1 positive-recurrent solutions to the Hofstadter $Q$-recurrence, and it finds two families (one family modulo shifting) of period 2 solutions. These solutions consist of one constant sequence interleaved with one standard linear sequence. For example, the initial condition $\langle 2, 2 \rangle$ gives rise to the sequence $2, 2, 4, 2, 6, 2, 8, 2, \ldots$ [31, A275365]. There are 12 period 3 solution families (4 modulo shifting). One of these families includes Golomb's solution, and another includes Ruskey's sequence. The other two families consist of eventually quasilinear solutions with two constant sequences and one standard linear sequence (and appear to be related to each other). One of these families includes sequence A264756 [31]; the other includes sequences A283878 and A284429 [31]; There are 12 period 4 families (5 modulo shifting), all of which are quasilinear with constant and standard linear sequences. Some of these families include the period 2 solutions, but each such family also include additional solutions that do not have period 2. There are 35 period 5 families (7 modulo shifting). Again, all of these are quasilinear. But, one of these families has a steep linear subsequence [31, A269328]. There is a lot more variety beginning at period 6. There are 294 solution families (86 modulo shifting) in this case. These solutions include quadratics as well as mixing of exponentials with steep linears. A similar diversity of solutions exists with period 7, where there are 588 solution families (84 modulo shifting). Period 8 has at least 3256 families (610 modulo shifting). This is possibly all of them, but computing these caused Maple to report "Warning, limiting number of iterations reached." So, there may be more families here that the Maple implementation was unable to find. Similarly, period 9 has at least 15273 families (2279 modulo shifting). A file containing

| | | Index | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Initial Cond. | $m$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| $\langle 2, 2\rangle$ | 2 | 2 | 2 | 4 | 2 | 6 | 2 | 8 | 2 | 10 | 2 | 12 | 2 | 14 | 2 | 16 | 2 | 18 | 2 | 20 | 2 |
| $\langle 1, 0, 3, 3, 2\rangle$ | 3 | **1** | **0** | 3 | 3 | 2 | 6 | 3 | 2 | 9 | 3 | 2 | 12 | 3 | 2 | 15 | 3 | 2 | 18 | 3 | 2 |
| $\langle 0, 2, 3, 1\rangle$ | 3 | **0** | **2** | 3 | 1 | 3 | 6 | 1 | 3 | 9 | 1 | 3 | 12 | 1 | 3 | 15 | 1 | 3 | 18 | 1 | 3 |
| $\langle 2, 1\rangle$ | 3 | 2 | 1 | 3 | 5 | 1 | 3 | 8 | 1 | 3 | 11 | 1 | 3 | 14 | 1 | 3 | 17 | 1 | 3 | 20 | 1 |
| $\langle 4, 1, 0, 3, 3, 1, 4, 8\rangle$ | 4 | **4** | 1 | 0 | **3** | 3 | 1 | 4 | 8 | 7 | 1 | 4 | 12 | 11 | 1 | 4 | 16 | 15 | 1 | 4 | 20 |
| $\langle 5, 2, 0, 3, 6, 5, 2\rangle$ | 5 | 5 | 2 | 0 | **3** | 6 | 5 | 2 | 5 | 5 | 12 | 5 | 2 | 10 | 5 | 18 | 5 | 2 | 15 | 5 | 24 |
| $\langle 4, 0, 5, 6, 2, 6, 6, 3\rangle$ | 6 | **4** | **0** | 5 | 6 | 2 | 6 | 6 | 3 | 11 | 6 | 2 | 12 | 6 | 3 | 23 | 6 | 2 | 18 | 6 | 3 |
| $\langle 12, 6, 4, 6, 1, 6, 12, 10, 4\rangle$ | 6 | 12 | 6 | 4 | 6 | 1 | 6 | 12 | 10 | 4 | 6 | 13 | 6 | 12 | 16 | 4 | 6 | 25 | 6 | 12 | 26 |
| $\langle 7, 0, 8, 7, 7, 8, 4\rangle$ | 7 | 7 | 0 | 8 | 7 | 7 | 8 | 4 | 7 | 7 | 16 | 7 | 7 | 16 | 4 | 7 | 14 | 24 | 7 | 7 | 32 |

Table 3.2: Selected Solutions to the Hofstadter $Q$-recurrence. Bold terms violate eventual patterns.

information on all of the solution families examined (through period 9), modulo shifting, can be found at `http://github.com/nhf216/thesis/hof_small_periods.txt`. (The notation in this file is somewhat different from the notation in this chapter, and this is the notation used in the Maple package.) See Table 3.2 for the first 20 terms of some of these sequences (A275365, A264756, A283878, A284429, A283879, A269328, A264757, A283880, A283881).

### 3.3.2 Other Recurrences

We have already seen one example of a solution to a recurrence other than Hofstadter's (our example in 3.2.4). Our algorithm has also been used to examine what sorts of recurrences can be satisfied by exponential subsequences that appear. This has led to the observation that any homogeneous linear recurrence with positive coefficients that sum to at least 2 can be realized as a component sequence of some positive-recurrent solution to some Hofstadter-like recurrence [11]. See Chapter 5 for details of this construction.

In addition, we have found positive-recurrent solutions to other recurrences, including the Conolly recurrence [31, A275363] and the Hofstadter-Conway recurrence [31, A052928]. This second case is notable because the solution has period 2 with both subsequences linear. This can happen because the Hofstadter-Conway recurrence is not basic (see 3.2.5). As a result, the congruence classes of the constant terms in the linear polynomials end up determining much of the behavior.

We conclude with one more finding: a quasilinear solution including linear subsequences with slope strictly between 0 and 1 to a linear Hofstadter-like recurrence [31, A283904]. In the analysis of our algorithm, we specifically avoid discovering such solutions. But, in the Maple package, the user can specify values for the parameters $A_r$ to bypass the exhaustive search the algorithm performs, and the values specified can be strictly between 0 and $m$ (which the algorithm would not otherwise try). The implementation may then find solutions with such subsequences.

**Proposition 3.7.** *The initial condition* $\langle 1, 1\rangle$ *to the recurrence* $H(n) = H(n-2H(n-1)) + H(n-3H(n-2))$ *generates an eventually quasilinear sequence given by*

$$\begin{cases} H(6k) = 1 \\ H(6k+1) = 3k-1 \\ H(6k+2) = 3k+1 \\ H(6k+3) = 1 \\ H(6k+4) = 3k+1 \\ H(6k+5) = 3k+2 \end{cases}$$

*when the index is at least* 41*.*

*Proof.* As usual, the proof is by induction on the index. Generating the first 59 terms of the sequence verifies that the proposition is true through $k = 9$. So, suppose $k \geq 10$, and suppose the proposition holds for all smaller $k$ values. There are six cases to check. We check two of them; the rest are similar and are left as exercises.

$n = 6k$**:** We have

$$\begin{aligned} H(6k) &= H(6k - 2H(6k-1)) + H(6k - 3H(6k-2)) \\ &= H(6k - 2(3(k-1)+2)) + H(6k - 3(3(k-1)+1)) \\ &= H(2) + H(-3k+6) \\ &= 1, \end{aligned}$$

as required.

$n = 6k + 1$**:** We have

$$\begin{aligned} H(6k+1) &= H(6k+1-2H(6k)) + H(6k+1-3H(6k-1)) \\ &= H(6k+1-2\cdot 1) + H(6k+1-3\,(3\,(k-1)+2)) \\ &= H(6k-1) + H(-3k+3) \\ &= 3\,(k-1)+2 \\ &= 3k-1, \end{aligned}$$

as required.

$\square$

# Chapter 4

# Embedding Polynomials in Solutions to the Hofstadter $Q$-Recurrence

In Chapter 3, we saw an example of a Hofstadter-like recurrence with a solution including an equally-spaced quadratic subsequence. In this chapter, we describe how to obtain such a solution to the Hofstadter $Q$-recurrence itself. In fact, we construct eventually-quasipolynomial solutions to the $Q$-recurrence of all positive degrees. The algorithm in Chapter 3 was key to the original exploration leading to the main result in this chapter.

First, we define the following:

**Definition 4.1.** *Fix integers $d \geq 1$ and $k \geq -1$. Let*

$$p_{d,k}(n) = 3d\binom{n+k}{1+k} + \sum_{i=1}^{k}(3i+2)\binom{n-1+k-i}{k-i}.$$

Observe that $p_{d,k}$ is a polynomial in $n$ of degree $k+1$. In particular, $p_{d,-1} = 3d$, and $p_{d,0} = 3dn$. We will prove the following theorem:

**Theorem 4.2.** *Fix a degree $d \geq 1$. Define a sequence $(a_m)_{m\geq 1}$ as follows:*

$$a_{3dn+r} = \begin{cases} 3d-2 & 3dn+r=1 \\ 0 & 3dn+r=2 \\ p_{d,\frac{r}{3}}(n) & r \equiv 0 \pmod 3 \\ 3d & r \equiv 1 \pmod 3 \ \textit{ and } 3dn+r>2 \\ 3 & r \equiv 2 \pmod 3 \ \textit{ and } r \neq 3d-1 \ \textit{ and } 3dn+r>2 \\ 2 & r = 3d-1 \ \textit{ and } 3dn+r>2, \end{cases}$$

*where $0 \leq r < 3d$ always. Then, $(a_m)$ satisfies the Hofstadter $Q$-recurrence after an initial condition of length $3d+2$.*

We use the following lemmas:

**Lemma 4.3.** *For all integers $d \geq 1$ and $k \geq 0$ we have $p_{d,k}(n) = p_{d,k-1}(n) + p_{d,k}(n-1)$.*

*Proof.* We have

$$\begin{aligned}
p_{d,k-1}(n) + p_{d,k}(n-1) &= 3d\binom{n+k-1}{k} + \sum_{i=1}^{k-1}(3i+2)\binom{n-2+k-i}{k-i-1} \\
&\quad + 3d\binom{n+k-1}{1+k} + \sum_{i=1}^{k}(3i+2)\binom{n-2+k-i}{k-i} \\
&= 3d\left(\binom{n+k-1}{k} + \binom{n+k-1}{1+k}\right) \\
&\quad + \sum_{i=1}^{k-1}(3i+2)\left(\binom{n-2+k-i}{k-i-1} + \binom{n-2+k-i}{k-i}\right) \\
&\quad + (3k+2)\binom{n-2}{0}.
\end{aligned}$$

Applying Pascal's Identity yields

$$\begin{aligned}
p_{d,k-1}(n) + p_{d,k}(n-1) &= 3d\binom{n+k}{1+k} + \sum_{i=1}^{k-1}(3i+2)\binom{n-1+k-i}{k-i} + (3k+2) \\
&= 3d\binom{n+k}{1+k} + \sum_{i=1}^{k}(3i+2)\binom{n-1+k-i}{k-i} \\
&= p_{d,k}(n),
\end{aligned}$$

as required. ∎

**Lemma 4.4.** *For all integers $d \geq 1$, $k \geq 1$, and $n \geq 0$ we have*

$$p_{d,k}(n) \geq 3dn + 3k + 2.$$

*Proof.* First, we observe that

$$p_{d,k}(0) = 3d\binom{k}{1+k} + \sum_{i=1}^{k}(3i+2)\binom{k-i-1}{k-i}.$$

All of these binomial coefficients are zero, except when $i = k$, since $\binom{-1}{0} = 1$. So, $p_{d,k}(0) = 3k + 2$. This equals $3dn + 3k + 2$, and hence is greater than or equal to it, as required.

Now,

$$\begin{aligned}
p_{d,k}(1) &= 3d\binom{1+k}{1+k} + \sum_{i=1}^{k}(3i+2)\binom{k-i}{k-i} \\
&= 3d + \sum_{i=1}^{k}(3i+2) \\
&= 3d + 3\left(\frac{k^2+k}{2}\right) + 2k \\
&= \frac{3}{2}k^2 + \frac{7}{2}k + 3d.
\end{aligned}$$

So,

$$\begin{aligned}
p_{d,k}(1) - 3d \cdot 1 + 3k + 2 &= \frac{3}{2}k^2 + \frac{7}{2}k + 3d - 3d - 3k - 2 \\
&= \frac{3}{2}k^2 + \frac{1}{2}k - 2 \\
&= \frac{(3k+4)(k-1)}{2}.
\end{aligned}$$

This is greater than or equal to 0, since $k \geq 1$. So, $p_{d,k}(1) \geq 3d + 3k + 2$, as required.

Now, observe that $p_{d,k}$ has nonnegative coefficients, so it is convex. We have seen that its average slope on the interval $[0,1]$ is at least $3d$, so its derivative for $n > 1$ must be strictly larger than $3d$ everywhere. Therefore, since $p_{d,k}(1) \geq 3d + 3k + 2$, we can conclude that $p_{d,k}(n) \geq 3dn + 3k + 2$ for all $n \geq 0$. □

We will now prove Theorem 4.2. In this proof, we will use the notation $Q_P(n)$ to denote the $n^{th}$ term of the sequence we are constructing. When doing so, we are inductively assuming that the prior terms satisfy the $Q$-recurrence.

*Proof.* We will check the three congruence classes mod 3 separately for $m > 3d+2$. As usual, $m = 3dn + r$ for $0 \leq r < 3d$. We will proceed by induction, so in each case we will assume that all previous values of the sequence are what they should be. Also, in all cases, since $m > 3d + 2$, $m - 3d > 2$. (This will come up when deciding whether or not we are in the special initial conditions for the first two values.)

**$r \equiv 0 \pmod 3$:** Assume $r \equiv 0 \pmod 3$. Then, $m = 3dn + r$ for some $n$. For convenience, let $\ell = \frac{r}{3}$. We wish to show that $Q_P(3dn + r) = p_{d,\ell}(n)$. Let $c = 2$ if $r = 0$;

otherwise, let $c = 3$. We have,

$$\begin{aligned} Q_P(3dn+r) &= Q_P(3dn+r-Q_P(3dn+r-1)) \\ &\quad + Q_P(3dn+r-Q_P(3dn+r-2)) \\ &= Q_P(3dn+r-c) + Q_P(3dn+r-3d) \\ &= Q_P(3dn+r-c) + Q_P(3d(n-1)+r) \\ &= Q_P(3dn+r-c) + p_{d,\ell}(n-1). \end{aligned}$$

If $r = 0$, then $\ell = 0$ and

$$Q_P(3dn+r-c) = Q_P(3dn+r-2) = 3d = p_{d,\ell-1}(n).$$

If $r \neq 0$, then $\ell > 0$ and

$$Q_P(3dn+r-c) = Q_P(3dn+r-3) = p_{d,\ell-1}(n).$$

In either case, we have

$$Q_P(3dn+r) = p_{d,\ell-1}(n) + p_{d,\ell}(n-1).$$

By Lemma 4.3, this equals $p_{d,\ell}(n)$, as required.

$r \equiv 1 \pmod 3$**:** Assume $r \equiv 1 \pmod 3$. Then, $m = 3dn+r$ for some $n$. We wish to show that $Q_P(3dn+r) = 3d$. For convenience, let $\ell = \frac{r-1}{3}$. We have,

$$\begin{aligned} Q_P(3dn+r) &= Q_P(3dn+r-Q_P(3dn+r-1)) \\ &\quad + Q_P(3dn+r-Q_P(3dn+r-2)) \\ &= Q_P(3dn+r-p_{d,\ell}(n)) + Q_P(3dn+r-Q_P(3dn+r-2)). \end{aligned}$$

If $\ell = 0$, then $p_{d,\ell}(n) = 3dn$ and $r = 1$. So, in that case, $3dn+r-p_{d,\ell}(n) = r = 1$. Also, in that case $Q_P(3dn+r-2) = 2$, so

$$Q_P(3dn+r-Q_P(3dn+r-2)) = Q_P(3dn+r-2) = 2.$$

Since $Q_P(1) = 3d-2$, we obtain $Q_P(3dn+r) = 3d-2+2 = 3d$ in the case when $r = 1$.

Otherwise, we have $\ell \geq 1$. In that case, $p_{d,\ell}(n) \geq 3dn+3\ell+2$ by Lemma 4.4. But, $3\ell+2 = r+1$ so, $3dn+r-p_{d-1}(n) \leq -1$. This causes the first term to underflow, so $Q_P(3dn+r-p_{d,\ell}(n)) = 0$. Hence, $Q_P(3dn+r) = Q_P(3dn+r-Q_P(3dn+r-2))$. In this case, we know $r \neq 1$, so $Q_P(3dn+r-2) = 3$. This means that

$$Q_P(3dn+r-Q_P(3dn+r-2)) = Q_P(3dn+r-3) = 3d.$$

So, $Q_P(3dn+r) = 3d$, as required.

$r \equiv 2 \pmod 3$**:** Assume $r \equiv 2 \pmod 3$. Then, $m = 3dn + r$ for some $n$. Let $c = 2$ if $r = 3d-1$; otherwise, let $c = 3$. We wish to show that $Q_P(3dn+r) = c$. For convenience, let $\ell = \frac{r-2}{3}$. We have,

$$\begin{aligned}
Q_P(3dn+r) &= Q_P(3dn+r-Q_P(3dn+r-1)) \\
&\quad + Q_P(3dn+r-Q_P(3dn+r-2)) \\
&= Q_P(3dn+r-3d) + Q_P(3dn+r-p_{d,\ell}(n)) \\
&= Q_P(3d(n-1)+r) + Q_P(3dn+r-p_{d,\ell}(n)) \\
&= c + Q_P(3dn+r-p_{d,\ell}(n)).
\end{aligned}$$

If $\ell = 0$, then $p_{d,\ell}(n) = 3dn$ and $r = 2$. So, in that case, $3dn+r-p_{d,\ell}(n) = r = 2$. Since $Q_P(2) = 0$, we obtain $Q_P(3dn+r) = c$ in the case when $r = 2$.

Otherwise, we have $\ell \geq 1$. In that case, $p_{d,\ell}(n) \geq 3dn+3\ell+2$ by Lemma 4.4. But, $3\ell+2 = r$ so, $3dn+r-p_{d-1}(n) \leq 0$, an underflow in the second term. This implies that $Q_P(3dn+r-p_{d,\ell}(n)) = 0$, so $Q_P(3dn+r) = c$, as required.

$\square$

Note that the only place we use the $3i+2$ in the definition of $p_{d,k}(n)$ is to obtain the lower bound of $r+2$ on the polynomials that we needed when proving Theorem 4.2. So, $3i+2$ could be replaced by any larger expression, and the proof would still go through. Also, observe that this construction is not a direct generalization of Golomb's construction (see Section 2.2), as the $d = 1$ case has two constant pieces and one linear piece, unlike Golomb's, which has one constant piece and two linear pieces. Also, Golomb's sequence is *purely* quasilinear, whereas our $d = 1$ example is only eventually quasilinear. It is unknown whether there exist purely quasipolynomial solutions to the Hofstadter $Q$-recurrence of degrees greater than 1. As an example, we construct a solution to Hofstadter's recurrence with a cubic subsequence. To do this, we set $d = 3$, which means that the sequence values will depend on the congruence class mod 9 of the index. We observe that

$$\begin{aligned}
p_{3,0} &= 9n \\
p_{3,1} &= 9\binom{n+1}{2} + 5\binom{n-1}{0} = \frac{9}{2}n\left(n+1\right) + 5 \\
&= \frac{9}{2}n^2 + \frac{9}{2}n + 5 \\
p_{3,2} &= 9\binom{n+2}{3} + 5\binom{n}{1} + 8\binom{n-1}{0} = \frac{9}{6}n\left(n+1\right)\left(n+2\right) + 5n + 8 \\
&= \frac{3}{2}n^3 + \frac{9}{2}n^2 + 8n + 8.
\end{aligned}$$

So, our sequence is defined by $a_1 = 7$, $a_2 = 0$, and for $9n + r > 2$,

$$a_{9n+r} = \begin{cases} 9n & r = 0 \\ 9 & r = 1 \\ 3 & r = 2 \\ \frac{9}{2}n^2 + \frac{9}{2}n + 5 & r = 3 \\ 9 & r = 4 \\ 3 & r = 5 \\ \frac{3}{2}n^3 + \frac{9}{2}n^2 + 8n + 8 & r = 6 \\ 9 & r = 7 \\ 2 & r = 8. \end{cases}$$

After the initial condition $\langle 7, 0, 5, 9, 3, 8, 9, 2, 9, 9, 3\rangle$, repeated applications of the Hofstadter $Q$-recurrence produce the sequence [31, A264758]

$$\begin{gathered}
7, 0, 5, 9, 3, 8, 9, 2, 9, 9, 3, 14, 9, 3, 22, 9, 2, 18, 9, 3, 32, 9, 3, 54, \\
9, 2, 27, 9, 3, 59, 9, 3, 113, 9, 2, \ldots
\end{gathered}$$

# Chapter 5

# Embedding Linear-Recurrent Sequences in Solutions to Nested Recurrences

At the end of the paper where Ruskey [30] presents his solution to the Hofstadter $Q$-recurrence, he asks whether every linear recurrent sequence exists as an equally-spaced subsequence of a solution to some meta-Fibonacci recurrence. Strictly-speaking, a meta-Fibonacci recurrence is a *two-term* basic linear nested recurrence [8]. In this chapter, we answer Ruskey's question positively for (general) basic linear recurrences and linear recurrences with positive coefficients. As was the case in Chapter 4, the algorithm in Chapter 3 was key to the original exploration leading to the main result in this chapter.

In particular, our proof is constructive. Our main theorem is the following:

**Theorem 5.1.** *Let $(a(n))_{n\geq 0}$ be a sequence of positive integers satisfying the recurrence*

$$a(n) = \sum_{i=1}^{k} b_i a(n-i),$$

*for some positive integer $k$ and nonnegative integers $b_1, b_2, \ldots, b_k$ whose sum is at least 2. Then, there is a sequence $(q(n))_{n\geq 0}$ satisfying some basic linear nested recurrence such that $q(2kn) = a(n)$ for all $n \geq 0$. (We call the number $2k$ the* quasi-period *of the sequence $(q(n))$.)*

*Proof.* Let $(a(n))_{n\geq 0}$ be a sequence of positive integers satisfying the recurrence

$$a(n) = \sum_{i=1}^{k} b_i a(n-i),$$

for some positive integer $k$ and nonnegative integers $b_1, b_2, \ldots, b_k$ whose sum is at least 2. For each $r$ from 0 to $k-1$, define the sequence $\left(a^{(r)}(n)\right)_{n\geq 1}$ as

$$a^{(r)}(n) = \sum_{i=1}^{r} b_i a^{(r)}(n-k-i+r) + \sum_{i=r+1}^{k} b_i a^{(r)}(n-i+r)$$

with $a^{(r)}(i) = a(i)$ for $i \leq k$. Notice that $a^{(0)}(n) = a(n)$, and the other sequences satisfy similar recurrences with the coefficients cycled. Since the coefficients of the recurrences are nonnegative and sum to at least 2, the sequences $\left(a^{(r)}(n)\right)_{n\geq 1}$ exhibit superlinear growth for all $r$.

Now, for all $m$, define the sequence $(q(n))_{n\geq 0}$ as follows:

$$\begin{cases} q(2mk+2j) = a^{(j)}(m) & 0 \leq j < k \\ q(2mk+2j+1) = 2k(k-j) & 0 \leq j < k, \end{cases}$$

We claim that $(q(n))_{n\geq 0}$, with the extension $q(n) = 0$ for $n < 0$, eventually satisfies the basic linear nested recurrence

$$M_a(n) = M_a(n - M_a(n-2)) + \sum_{i=1}^{k} b_i M_a(n - M_a(n-(2i-1))).$$

Notice that this would imply the desired result, since the quasi-period will be $2k$. Let $h$ be an integer satisfying all of the following constraints:

- $h \geq 2k - 1$
- $h \geq 2$
- For all $r$, whenever $m \geq h$, $a^{(r)}(m-1) \geq 2(m+1)k$.

We define a function $L$ as follows:

$$L(n) = \begin{cases} q(n) & n \leq h \\ L(n - L(n-2)) + \displaystyle\sum_{i=1}^{k} b_i L(n - L(n-(2i-1))) & n > h. \end{cases}$$

In other words, $L$ eventually satisfies the recurrence $M_a$, and it has an initial condition of length $h$ that matches $(q(n))$. The first two conditions on $h$ are required to make $L$ well-defined. Since all the linear recurrent sequences under consideration grow superlinearly, the third condition is satisfied by all sufficiently large numbers. Hence, such an $h$ exists, and all larger values would also be valid choices for $h$.

We wish to show that $L(n) = q(n)$ for all $n$. We will proceed by induction on $n$. The base case is covered by the fact that $L(n)$ is defined to equal $q(n)$ for $n \leq h$. So, we will show that, for $n > h$, $L(n) = q(n)$ under the assumption that $L(p) = q(p)$ for all $1 \le p < n$. For this, we will split into two cases:

$n$ **is odd:** Since $n$ is odd, it is of the form $2mk + 2j + 1$ for some $m \ge 0$ and some $0 \le j < k$. By our choice of $h$, $a^{(r)}(m-1) \ge 2\,(m+1)\,k$ and $a^{(r)}(m) \ge 2\,(m+2)\,k$. In particular, both of these are greater than $2mk + 2j + 1$. Using this fact, we have

$$\begin{aligned}
L(n) &= L(n - L(n-2)) + \sum_{i=1}^{k} b_i L(n - L(n - (2i-1))) \\
&= L(n - q(n-2)) + \sum_{i=1}^{k} b_i L(n - q(n - (2i-1))) \\
&= L(2mk + 2j + 1 - q(2mk + 2j - 1)) \\
&\quad + \sum_{i=1}^{k} b_i L(2mk + 2j + 1 - q(2mk + 2j + 1 - (2i-1))) \\
&= L(2mk + 2j + 1 - q(2mk + 2\,(j-1) + 1)) \\
&\quad + \sum_{i=1}^{k} b_i L(2mk + 2j + 1 - q(2mk + 2\,(j - i + 1))) \\
&= L(2mk + 2j + 1 - q(2mk + 2\,(j-1) + 1)) \\
&\quad + \sum_{i=1}^{j+1} b_i L(2mk + 2j + 1 - a^{(j-i+1)}(m)) \\
&\quad + \sum_{i=j+2}^{k} b_i L(2mk + 2j + 1 - a^{(k+j-i+1)}(m-1)) \\
&= L(2mk + 2j + 1 - q(2mk + 2\,(j-1) + 1)) + \sum_{i=1}^{j+1} b_i \cdot 0 + \sum_{i=j+2}^{k} b_i \cdot 0.
\end{aligned}$$

If $j = 0$, then $q(2mk + 2\,(j-1) + 1) = 2k$; otherwise, $q(2mk + 2\,(j-1) + 1) = 2k(k - j + 1)$. In both cases, it is of the form $2ks$ for some $s$. So, we have

$$\begin{aligned}
L(n) &= L(2mk + 2j + 1 - 2ks) \\
&= L(2(m-s)k + 2j + 1) \\
&= 2k\,(k - j) \\
&= q(2mk + 2j + 1) \\
&= q(n),
\end{aligned}$$

as required.

$n$ **is even:** Since $n$ is even, it is of the form $2mk + 2j$ for some $m \geq 0$ and some $0 \leq j < k$. By our choice of $h$, $a^{(r)}(m-1) \geq 2\,(m+1)\,k$ and $a^{(r)}(m) \geq 2\,(m+2)\,k$. In particular, both of these are greater than $2mk + 2j$. Using this fact, we have

$$\begin{aligned}
L(n) &= L(n - L(n-2)) + \sum_{i=1}^{k} b_i L(n - L(n - (2i-1))) \\
&= L(n - q(n-2)) + \sum_{i=1}^{k} b_i L(n - q(n - (2i-1))) \\
&= L(2mk + 2j - q(2mk + 2j - 2)) \\
&\quad + \sum_{i=1}^{k} b_i L(2mk + 2j - q(2mk + 2j - (2i-1))) \\
&= L(2mk + 2j - q(2mk + 2\,(j-1))) \\
&\quad + \sum_{i=1}^{k} b_i L(2mk + 2j - q(2mk + 2\,(j-i) + 1))
\end{aligned}$$

If $j = 0$, then we will have $q(2mk + 2\,(j-1)) = a^{(k-1)}(m-1)$. Otherwise, we will have $q(2mk + 2\,(j-1)) = a^{(j-1)}(m)$. In either case, we have $L(2mk + 2j - q(2mk + 2\,(j-1))) = 0$. So,

$$\begin{aligned}
L(n) &= 0 + \sum_{i=1}^{k} b_i L(2mk + 2j + 1 - q(2mk + 2\,(j-i) + 1)) \\
&= \sum_{i=1}^{j} b_i L(2mk + 2j - 2k(k - (j-i))) \\
&\quad + \sum_{i=j+1}^{k} b_i L(2mk + 2j - 2k(k - (k + j - i))) \\
&= \sum_{i=1}^{j} b_i L(2mk + 2j - 2k\,(k - j + i)) + \sum_{i=j+1}^{k} b_i L(2mk + 2j - 2k\,(i-j)) \\
&= \sum_{i=1}^{j} b_i L(2\,(m - k - i + j)\,k + 2j) + \sum_{i=j+1}^{k} b_i L(2\,(m - i + j)\,k + 2j) \\
&= \sum_{i=1}^{j} b_i q(2\,(m - k - i + j)\,k + 2j) + \sum_{i=j+1}^{k} b_i q(2\,(m - i + j)\,k + 2j) \\
&= \sum_{i=1}^{j} b_i a^{(j)}(m - k - i + j) + \sum_{i=j+1}^{k} b_i a^{(j)}(m - i + j)
\end{aligned}$$

$$= a^{(j)}(m)$$

$$= q(2mk + 2j)$$

$$= q(n),$$

as required. □

In the case where $(a(n))_{n\geq 0}$ is the Fibonacci sequence starting from 5, this construction does not give Ruskey's sequence. Rather, we obtain the sequence

$$\langle 5, 8, 5, 4, 8, 8, 8, 4, 13, 8, 13, 4, 21, 8, 21, 4, \ldots \rangle$$

that eventually satisfies the recurrence

$$M_a(n) = M_a(n - M_a(n-1)) + M_a(n - M_a(n-2)) + M_a(n - M_a(n-3)).$$

The Fibonacci numbers each appear twice in this sequence because the Fibonacci recurrence is invariant under rotation (and each rotation of it appears once).

Since any linear recurrent sequence satisfies infinitely many linear recurrences, this construction actually gives infinitely many nested recurrences that can include a given linear-recurrent sequence. In addition, the construction can be tweaked in a number of ways to yield slightly different sequences. For example, one could start from a rotation of the desired sequence. Or, the initial conditions for the rotations could be chosen differently, since their values are not critical to the construction. (We only care about the growth rate and recurrent behavior of the rotations.) But, none of these modifications would suffice to cause our construction to yield Ruskey's sequence, since his sequence has quasi-period 3 and our construction only yields sequences with even quasi-periods. This fact seems to indicate that there are many more nested recurrences generating a given linear recurrent sequence than our construction can generate.

In our construction, we put two constraints on the $b$ values. First, we require them to be nonnegative. With our conventions, it would be impossible to have a solution to a generalized meta-Fibonacci recurrence with infinitely many nonpositive entries. There are many linear recurrent sequences with positive terms but some negative coefficients. But, our construction fails for these sequences, since some rotation of such a sequence will have infinitely many nonpositive entries. Ruskey's question remains open for such sequences. Second, we require the sum of the $b$ values to be at least 2. This was necessary to force the terms of $(a(n))$ to grow superlinearly. If the $b$ values sum to zero, then they must all be zero, in which case the sequence $(a(n))$ is eventually zero, and, hence, not a sequence of positive integers. If the $b$ values sum to 1, then all of them must be zero except for one. So, the recurrence we obtain is $a(n) = b_i a(n-i)$ for some $i$. So, in this case, $(a(n))$ is eventually periodic. Eventually constant sequences eventually satisfy the recurrence $M(n) = M(n - M(n-1))$, but it is unclear whether higher periods can always be realized within generalized meta-Fibonacci sequences.

The following example should illustrate most of the nuances of our construction. Consider the sequence $(a(n))_{n\geq 0}$ defined by $a(0) = 30$, $a(1) = 40$, $a(2) = 60$, and $a(n) = a(n-1) + 2a(n-3)$ for $n \geq 3$. (The large initial values allow us to avoid having an unreasonably long initial condition.) The first few terms of this sequence are $(30, 40, 60, 120, 200, 320, 560, 960, \ldots)$. The rotations of $(a(n))$ have the same initial conditions and are given by the following recurrences:

$$\begin{aligned}
a^{(0)}(n) &= a^{(0)}(n-1) + 2a^{(0)}(n-3) \quad (30, 40, 60, 120, 200, 320, 560, 960, \ldots)\\
a^{(1)}(n) &= a^{(1)}(n-3) + 2a^{(1)}(n-2) \quad (30, 40, 60, 110, 160, 280, 430, 720, \ldots)\\
a^{(2)}(n) &= a^{(2)}(n-2) + 2a^{(2)}(n-1) \quad (30, 40, 60, 160, 380, 920, 2220, 5360, \ldots)
\end{aligned}$$

The construction gives the sequence $(q(n))_{n\geq 0}$ defined by

$$\begin{cases} q(6m+2j) = a^{(j)}(m) & 0 \leq j < 4 \\ q(6m+2j+1) = 6(3-j) & 0 \leq j < 4 \end{cases}$$

as eventually satisfying the recurrence

$$M_a(n) = M_a(n - M_a(n-1)) + M_a(n - M_a(n-2)) + 2M_a(n - M_a(n-5)).$$

Sure enough, the initial condition

$$\langle 30, 18, 30, 12, 30, 6, 40, 18, 40, 12, 40, 6, 60, 18, 60, 12, 60, 6\rangle$$

suffices. The next term is

$$\begin{aligned}
M_a(18) &= M_a(18 - M_a(17)) + M_a(18 - M_a(16)) + 2M_a(18 - M_a(13)) \\
&= M_a(18 - 6) + M_a(18 - 60) + 2M_a(18 - 18) \\
&= M_a(12) + M_a(-42) + 2M_a(0) \\
&= 60 + 0 + 2 \cdot 30 \\
&= 120 \\
&= a^{(0)}(3),
\end{aligned}$$

as required. The term after this is

$$\begin{aligned}
M_a(19) &= M_a(19 - M_a(18)) + M_a(19 - M_a(17)) + 2M_a(19 - M_a(14)) \\
&= M_a(19 - 120) + M_a(19 - 6) + 2M_a(19 - 60) \\
&= M_a(-101) + M_a(13) + 2M_a(-41) \\
&= 0 + 18 + 2 \cdot 0 \\
&= 18 \\
&= 6(3 - 0),
\end{aligned}$$

as required. The rest of the desired terms can continue to be generated this way.

# Chapter 6

# A Slow Solution to a Hofstadter-like Nested Recurrence

In this chapter, we present a new example of a slow solution to a nested recurrence. (See Section 2.3 in Chapter 2 for background on slow solutions.) Most of the known examples of slow sequences have at least one of the following properties (see p. 12):

- A "Conway-like term" (e.g. the Hofstadter-Conway recurrence [28]).
- A "shift" in at least one of the recurrence terms (e.g. Conolly's recurrence [7]).

In fact, the only previously known ones that have neither property are the $V$-sequence [3] and sequences constructed from it [20]. We decided to search for additional slow, Hofstadter-like sequences without shifts and without inner positive coefficients. The investigation of Hofstadter and Huber [19] (see p. 12) empirically rules out two-term recurrences, so we began our search by considering the generic 3-term recurrence

$$Q_{r,s,t}(n) = Q_{r,s,t}(n - Q_{r,s,t}(n-r)) + Q_{r,s,t}(n - Q_{r,s,t}(n-s)) + Q_{r,s,t}(n - Q_{r,s,t}(n-t))$$

with integers $0 < r < s < t$. The all-ones initial condition proved fruitless in our investigation. However, the initial conditions $\langle 1, 2, 3, 4 \rangle$ generate the $V$-sequence as well (offset by 3 terms) [3]. Thus, we focused our search on slow sequences with initial conditions of the form $Q_{r,s,t}(i) = i$ for $i \leq t$. This allowed us to find an interesting sequence with $(r, s, t) = (1, 2, 3)$. In this chapter, we prove that this sequence is slow. In fact, we completely characterize the terms of this sequence and exhibit an efficient algorithm for computing the $n^{th}$ term. In particular, each term of this sequence appears at most twice, in contrast to the $V$-sequence, whose terms appear at most three times [3].

Section 9.4 in Chapter 9 discusses failed attempts to generalize this construction.

We consider the sequence defined by the recurrence

$$B(n) = B(n - B(n-1)) + B(n - B(n-2)) + B(n - B(n-3))$$

and the initial conditions $\langle 1,2,3,4,5\rangle$. The first few terms of this sequence are [31, A278055]

$$1,2,3,4,5,6,6,7,8,9,9,10,11,12,12,13,14,15,15,16,17,17,18,18,19,20,21,21,\dots$$

The main thing we wish to prove is the following:

**Theorem 6.1.** *For all $n$, $B(n)-B(n-1)\in\{0,1\}$. In other words, the sequence $(B(n))_{n\geq 1}$ is slow.*

We actually prove considerably more than just Theorem 6.1. We completely determine the structure of this sequence. In the terms listed above, each positive integer appears no more than twice (and at least once). We show that this is the case for all numbers, and we completely characterize which numbers repeat.

We make use of the following auxiliary sequence $(a_i)_{i\geq 1}$. Let $a_1=3$, and for $i\geq 1$, let $a_i=3a_{i-1}-1$. This sequence [31, A057198] has the closed form $a_i=\frac{5}{2}3^{i-1}+\frac{1}{2}$. We have the following theorem.

**Theorem 6.2.** *Let $m$ be a positive integer. If there exists some integer $k\geq 1$ such that $m=k\cdot 3^i+a_i$ for some $i\geq 1$, then $m$ appears in the $B$-sequence twice. Otherwise, $m$ appears once. Furthermore, the $B$-sequence is monotone increasing.*

Theorem 6.2 implies Theorem 6.1, since Theorem 6.2 asserts both that the sequence is monotone and that each positive integer appears in the sequence. Throughout the rest of this section, we end up proving Theorem 6.2, and consequently Theorem 6.1, by induction. In doing so, we frequently assume that Theorem 6.2 holds up to some point. To make this clear, we define the following indexed families of propositions (where $m$ and $n$ are positive integers):

- Let $P_m$ denote the proposition "For all integers $1\leq m'\leq m$, if there exists some integer $k\geq 1$ such that $m'=k\cdot 3^i+a_i$ for some $i\geq 1$, then $m'$ appears in the $B$-sequence twice. Otherwise, $m'$ appears once. Furthermore, the $B$-sequence is monotone increasing as long as its terms are at most $m$." In this way, $P_m$ is essentially the statement "Theorem 6.2 holds through *value* $m$."

- Let $T_n$ denote the proposition "The first $n$ terms of the $B$-sequence are monotone increasing. Furthermore, for all $m$ appearing as one of these first $n$ terms, if there exists some integer $k \geq 1$ such that $m = k \cdot 3^i + a_i$ for some $i \geq 1$, then $m$ appears in these first $n$ terms twice (unless this second occurrence would be in position $n+1$). Otherwise, $m$ appears once." In this way, $T_n$ is essentially the statement "Theorem 6.2 holds through *index* $n$."

It should be clear from these definitions that the following are equivalent:

- Theorem 6.2 is true.
- $P_m$ holds for all $m \geq 1$.
- $T_n$ holds for all $n \geq 1$.

We call a pair of positive integers $(k, i)$ a *witness pair* for $m$ if $m = k \cdot 3^i + a_i$, and we call such an $i$ a *witness* for $m$. (Theorem 6.2 says that a value $m$ is repeated if and only if it has a witness.) We now show that every $m$ has at most one witness.

**Lemma 6.3.** *For any positive integer $m$, there is at most one $i \geq 1$ such that $m \equiv a_i \pmod{3^i}$.*

*Proof.* Suppose for a contradiction that, for some integers $i, j \geq 1$, $k_1 \cdot 3^i + a_i = k_2 \cdot 3^{i+j} + a_{i+j}$. Then

$$a_{i+j} - a_i = k_1 \cdot 3^i - k_2 \cdot 3^{i+j} = 3^i(k_1 + k_2 \cdot 3^j).$$

In particular, $a_{i+j} - a_i$ must be divisible by $3^i$.

But, using the closed form,

$$\begin{aligned} a_{i+j} - a_i &= \left(\frac{5}{2} \cdot 3^{i+j-1} + \frac{1}{2}\right) - \left(\frac{5}{2} \cdot 3^{i-1} + \frac{1}{2}\right) = \frac{5}{2}\left(3^{i+j-1} - 3^{i-1}\right) \\ &= \frac{5}{2} \cdot 3^{i-1}\left(3^j - 1\right). \end{aligned}$$

This is clearly not divisible by $3^i$, a contradiction. Therefore, no such $i$ and $j$ can exist, so there is at most one $i \geq 1$ such that $m \equiv a_i \pmod{3^i}$, as required. $\square$

For a value $m$, we now examine the number of values less than $m$ that are repeated. Define $r(m,i) = \max\left(0, \left\lfloor \frac{m-a_i-1}{3^i} \right\rfloor\right)$. This floored quantity counts the witness pairs $(k,i)$ for numbers less than $m$. If $P_m$ holds, then this is also the number of repeated values $m' < m$ with witness $i$. If we now let

$$r(m) = \sum_{i=1}^{\infty} r(m,i),$$

we have that $r(m)$ is the total number of repeated values less than $m$ (provided that $P_m$ holds.) This sum converges because only the logarithmically many terms with $a_i - 1 \leq m$ are nonzero.

We now have the following lemmas.

**Lemma 6.4.** *Let $m$ be a positive integer. Suppose $P_{m-1}$ holds. Then, $B(m+r(m)-1) = m-1$, and $B(m+r(m)) \geq m$. (In other words $m+r(m)-1$ is the last index in our sequence with value at most $m-1$.)*

*Proof.* The number of terms before the first occurrence of a term greater than or equal to $m$ is at least $m-1$, since each number smaller than $m$ must appear at least once. The first occurrence of such a term is "delayed" by 1 index for every smaller value that is repeated. The number of such repeated values is $r(m)$. So, there are $m-1+r(m)$ terms before the first occurrence of a term greater than or equal to $m$. This means that the last occurrence of $m-1$ is in position $m+r(m)-1$, as required. $\square$

An immediate consequence of Lemma 6.4 is that $B(m+r(m))$ in fact *equals* $m$, provided that $P_m$ holds.

**Lemma 6.5.** *Let $m$ be a multiple of 3. If $i \geq 2$ is a witness for $m-1$, then $r(m,i) = r\left(\frac{m}{3}, i-1\right)+1$. Otherwise, $r(m,i) = r\left(\frac{m}{3}, i-1\right)$.*

*Proof.* The lemma is clearly true if $a_i + 1 \geq m$, so we can assume without loss of generality that $a_i + 1 < m$ and thereby ignore the max in the definition of $r(m,i)$ when proving this lemma.

We have

$$r(m,i) = \left\lfloor \frac{m-a_i-1}{3^i} \right\rfloor = \left\lfloor \frac{m}{3^i} - \frac{a_i+1}{3^i} \right\rfloor$$

and

$$r\left(\frac{m}{3}, i-1\right) = \left\lfloor \frac{\frac{m}{3} - a_{i-1} - 1}{3^{i-1}} \right\rfloor = \left\lfloor \frac{m}{3^i} - \frac{a_{i-1}+1}{3^{i-1}} \right\rfloor.$$

Since $a_i = \frac{5}{2} \cdot 3^{i-1} + \frac{1}{2}$,

$$\frac{a_i + 1}{3^i} = \frac{5}{6} + \frac{1}{2 \cdot 3^i}$$

and

$$\frac{a_{i-1}+1}{3^{i-1}} = \frac{5}{6} + \frac{1}{2 \cdot 3^{i-1}}.$$

The first of these definitely smaller, so $r(m,i) \geq r\left(\frac{m}{3}, i-1\right)$. Furthermore, the above fractions differ by $\frac{1}{3^i}$, so $r(m,i) \leq r\left(\frac{m}{3}, i-1\right) + 1$.

The only way they could not be equal is if there is some integer $\ell$ such that

$$\frac{m}{3^i} - \frac{a_{i-1}+1}{3^{i-1}} < \ell \leq \frac{m}{3^i} - \frac{a_i+1}{3^i}.$$

Since the bounds differ by $\frac{1}{3^i}$ and they have common denominator $3^i$, this can only happen if $\ell = \frac{m}{3^i} - \frac{a_i+1}{3^i}$. This gives that $m - a_i + 1 = \ell \cdot 3^i$, or $m - 1 = \ell \cdot 3^i + a_i$ for some integer $\ell$. Since $a_i + 1 < m$, we must have $\ell \geq 1$. So, for $r(m,i) = r\left(\frac{m}{3}, i-1\right) + 1$, we obtain that $i$ must be a witness for $m-1$, as required. □

**Lemma 6.6.** *Let $m$ be a multiple of 3. Then,*

$$\frac{m}{3} + r\left(\frac{m}{3}\right) = \begin{cases} r(m) + 1, & \textit{if } m-1 \textit{ has a witness;} \\ r(m) + 2, & \textit{if } m-1 \textit{ does not have a witness.} \end{cases}$$

*Proof.* As a consequence of Lemma 6.5 and Lemma 6.3,

$$r(m) = \begin{cases} r(m,1) + r\left(\frac{m}{3}\right), & \text{if } m-1 \text{ does not have a witness;} \\ r(m,1) + r\left(\frac{m}{3}\right) + 1, & \text{if } m-1 \text{ has a witness.} \end{cases}$$

We also have

$$r(m,1) = \left\lfloor \frac{m - a_1 - 1}{3} \right\rfloor = \left\lfloor \frac{m-4}{3} \right\rfloor = \frac{m}{3} - 2.$$

Substituting this into the above and rearranging terms gives the required form. □

**Lemma 6.7.** *Let $m$ be a multiple of 3. Then $m-1$ has a witness if and only if $\frac{m}{3}$ has a witness.*

*Proof.* ($\Longrightarrow$) Suppose $m-1 = k\cdot 3^i + a_i$ for some positive integers $k$ and $i$. Then, $m = k\cdot 3^i + a_i + 1$. But, $a_i = 3a_{i-1} - 1$, so $m = k\cdot 3^i + 3a_{i-1}$. This means that $\frac{m}{3} = k\cdot 3^{i-1} + a_{i-1}$, so $i-1$ is a witness for $\frac{m}{3}$.

($\Longleftarrow$) Suppose $\frac{m}{3} = k\cdot 3^i + a_i$ for some positive integers $k$ and $i$. Then, $m = 3k\cdot 3^i + 3a_i$. But, $a_{i+1} = 3a_i - 1$, so $m = k\cdot 3^{i+1} + a_{i+1} + 1$. This means that $m-1 = k\cdot 3^{i+1} + a_{i+1}$, so $i+1$ is a witness for $m-1$.

□

**Lemma 6.8.** *Let $m \geq 6$ be a multiple of 3. Suppose $P_{m-1}$ holds. Then, if $m-1$ repeats we have $B(r(m)+1) = \frac{m}{3}$. If $m-1$ does not repeat we have $B(r(m)+1) = \frac{m}{3} - 1$. In both cases we have*

$$\begin{cases} B(r(m)+2) = \frac{m}{3}; \\ B(r(m)+3) = \frac{m}{3} + 1. \end{cases}$$

*Proof.* We examine the two cases separately.

$m-1$ **repeats:** Then, $m-1$ has a witness. So, by Lemma 6.6, $r(m)+1 = \frac{m}{3} + r\left(\frac{m}{3}\right)$. By Lemma 6.4, $B(r(m)+1) = \frac{m}{3}$. Furthermore, by Lemma 6.7, $\frac{m}{3}$ has a witness (and hence repeats), so $B(r(m)+2) = \frac{m}{3}$ as well. Since values appear at most twice, we then have $B(r(m)+3) = \frac{m}{3} + 1$, as required.

$m-1$ **does not repeat:** Then, $m-1$ does not have a witness. So, by Lemma 6.6, $r(m)+2 = \frac{m}{3} + r\left(\frac{m}{3}\right)$. By Lemma 6.4, $B(r(m)+2) = \frac{m}{3}$ and $B(r(m)+1) = \frac{m}{3} - 1$. Furthermore, by Lemma 6.7, $\frac{m}{3}$ has no witness (and hence does not repeat), so $B(r(m)+3) = \frac{m}{3} + 1$.

□

We are now ready to prove Theorem 6.2.

*Proof.* The proof is by induction on $n$, the index in the sequence. For the base case, observe that each term in the initial condition appears once, and no such term has a witness.

Now, suppose that $T_{n-1}$ holds, and suppose that we wish to show that $B(n) = m$ for some $m \geq 6$. Also, suppose that $P_{m-1}$ holds. There are seven cases to consider, which cover all possibilities. (Note that no repeated term is congruent to 1 mod 3, since $a_1$ is divisible by 3 and $a_i \equiv 2 \pmod 3$ for all $i \geq 2$.)

$m \equiv 0 \pmod 3$**, first occurrence,** $m-1$ **not repeated:** In this case, $m-1$ has no witness and, by Lemma 6.4, $n = m + r(m)$. We have (using Lemma 6.8)

$$\begin{aligned}
B(n) &= B(n - B(n-1)) + B(n - B(n-2)) + B(n - B(n-3)) \\
&= B(m + r(m) - (m-1)) + B(m + r(m) - (m-2)) \\
&\quad + B(m + r(m) - (m-3)) \\
&= B(r(m)+1) + B(r(m)+2) + B(r(m)+3) \\
&= \left(\frac{m}{3} - 1\right) + \frac{m}{3} + \left(\frac{m}{3} + 1\right) \\
&= m,
\end{aligned}$$

as required.

$m \equiv 0 \pmod 3$**, first occurrence,** $m-1$ **repeated:** In this case, $m-1$ has a witness and, by Lemma 6.4, $n = m + r(m)$. We have (using Lemma 6.8)

$$\begin{aligned}
B(n) &= B(n - B(n-1)) + B(n - B(n-2)) + B(n - B(n-3)) \\
&= B(m + r(m) - (m-1)) + B(m + r(m) - (m-1)) \\
&\quad + B(m + r(m) - (m-2)) \\
&= B(r(m)+1) + B(r(m)+1) + B(r(m)+2) \\
&= \frac{m}{3} + \frac{m}{3} + \frac{m}{3} \\
&= m,
\end{aligned}$$

as required.

$m \equiv 0 \pmod 3$**, second occurrence,** $m-1$ **not repeated:** In this case, $m-1$ has no witness and, by Lemma 6.4, $n = m + r(m) + 1$. We have (using Lemma 6.8)

$$B(n) = B(n - B(n-1)) + B(n - B(n-2)) + B(n - B(n-3))$$

$$
\begin{aligned}
&= B(m+r(m)+1-m) + B(m+r(m)+1-(m-1)) \\
&\quad + B(m+r(m)+1-(m-2)) \\
&= B(r(m)+1) + B(r(m)+2) + B(r(m)+3) \\
&= \left(\frac{m}{3}-1\right) + \frac{m}{3} + \left(\frac{m}{3}+1\right) \\
&= m,
\end{aligned}
$$

as required.

$m \equiv 0 \pmod 3$**, second occurrence,** $m-1$ **repeated:** In this case, $m-1$ has a witness and, by Lemma 6.4, $n = m + r(m) + 1$. We have (using Lemma 6.8)

$$
\begin{aligned}
B(n) &= B(n-B(n-1)) + B(n-B(n-2)) + B(n-B(n-3)) \\
&= B(m+r(m)+1-m) + B(m+r(m)+1-(m-1)) \\
&\quad + B(m+r(m)+1-(m-1)) \\
&= B(r(m)+1) + B(r(m)+1) + B(r(m)+2) \\
&= \frac{m}{3} + \frac{m}{3} + \frac{m}{3} \\
&= m,
\end{aligned}
$$

as required.

$m \equiv 1 \pmod 3$**:** In this case, $m-1$ is divisible by 3 and therefore definitely repeats (since $a_1 = 3$). This also means that $r(m-1) = r(m) - 1$. By Lemma 6.4, $n = m + r(m)$. We have (using Lemma 6.8)

$$
\begin{aligned}
B(n) &= B(n-B(n-1)) + B(n-B(n-2)) + B(n-B(n-3)) \\
&= B(m+r(m)-(m-1)) + B(m+r(m)-(m-1)) \\
&\quad + B(m+r(m)-(m-2)) \\
&= B(r(m)+1) + B(r(m)+1) + B(r(m)+2) \\
&= B(r(m-1)+2) + B(r(m-1)+2) + B(r(m-1)+3) \\
&= \frac{m-1}{3} + \frac{m-1}{3} + \left(\frac{m-1}{3}+1\right) \\
&= m,
\end{aligned}
$$

as required.

$m \equiv 2 \pmod 3$**, first occurrence:** In this case, $m-2$ is divisible by 3 and therefore definitely repeats. This also means that $r(m-2) = r(m) - 1$. By Lemma 6.4, $n = m + r(m)$. We have (using Lemma 6.8)

$$\begin{aligned}
B(n) &= B(n - B(n-1)) + B(n - B(n-2)) + B(n - B(n-3)) \\
&= B(m + r(m) - (m-1)) + B(m + r(m) - (m-2)) \\
&\quad + B(m + r(m) - (m-2)) \\
&= B(r(m)+1) + B(r(m)+2) + B(r(m)+2) \\
&= B(r(m-2)+2) + B(r(m-2)+3) + B(r(m-2)+3) \\
&= \frac{m-2}{3} + \left(\frac{m-2}{3} + 1\right) + \left(\frac{m-2}{3} + 1\right) \\
&= m,
\end{aligned}$$

as required.

$m \equiv 2 \pmod 3$**, second occurrence:** In this case, $m$ has a witness, so $r(m+1) = r(m) + 1$. Also, $r(m-2) = r(m) - 1$. By Lemma 6.4, $n = m + r(m) + 1$. We have (using Lemma 6.8)

$$\begin{aligned}
B(n) &= B(n - B(n-1)) + B(n - B(n-2)) + B(n - B(n-3)) \\
&= B(m + r(m) + 1 - m) + B(m + r(m) + 1 - (m-1)) \\
&\quad + B(m + r(m) + 1 - (m-2)) \\
&= B(r(m)+1) + B(r(m)+2) + B(r(m)+3) \\
&= B(r(m-2)+2) + B(r(m-2)+3) + B(r(m+1)+2) \\
&= \frac{m-2}{3} + \left(\frac{m-2}{3} + 1\right) + \left(\frac{m+1}{3}\right) \\
&= m,
\end{aligned}$$

as required.

$\square$

We have the following corollary.

**Corollary 6.9.** *We have*

$$\lim_{n\to\infty} \frac{B(n)}{n} = \frac{2}{3}.$$

*Proof.* If $B(n) = m$, then $n = m + r(m)$ or $n = m + r(m) + 1$. So, it will suffice to show that

$$\lim_{m\to\infty} \frac{m}{m + r(m)} = \frac{2}{3},$$

for which it is sufficient to show that

$$\lim_{m\to\infty} \frac{r(m)}{m} = \frac{1}{2}.$$

For each $i \geq 1$, we have

$$\lim_{m\to\infty} \frac{r(m,i)}{m} = \frac{1}{3^i}.$$

So,

$$\lim_{m\to\infty} \frac{r(m)}{m} = \lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{\infty} r(m,i) = \sum_{i=1}^{\infty} \lim_{m\to\infty} \frac{r(m,i)}{m} = \sum_{i=1}^{\infty} \frac{1}{3^i} = \frac{1}{2},$$

as required. □

Theorem 6.2 leads to an efficient algorithm for calculating $B(n)$. Observe that, for each $m$ and $i$, $r(m,i)$ can be computed efficiently. Since only logarithmically many terms in the sum for $r(m)$ are nonzero, this means that $r(m)$ can be computed efficiently.

To compute $B(n)$, we seek an $m$ such that $n = m + r(m)$. It may be the case that no such $m$ exists, in which case we need to be able to say that no such $m$ exists, and we need to find $m$ such that $n = m + r(m) + 1$. This task can be done efficiently using a binary search. We know that $B(n) \leq n$, so for an initial upper bound on $m$ we can use $n$ (and we can use 1 as a lower bound). So, in at most $O(\log(n))$ steps, we can either find an $m$ so that $n = m + r(m)$ or show that none exists. In the latter case, the final lower bound we find for $m$ is such that $n = m + r(m) + 1$. The total running time of this algorithm is $O(\log^2(n))$.

# Chapter 7

# Semi-Predictable Solutions

This chapter serves as a transition between the material in the preceding chapters and the material in the upcoming chapters. The former content is primarily concerned with determining initial conditions that produce prescribed sequence behavior; the latter content is primarily concerned with determining the behavior of sequences generated by prescribed initial conditions. This chapter's results contain elements of both groups. Proposition 7.2, the main result in this chapter, involves finding an initial condition to generate a specific type of sequence. But, it was originally discovered when exploring the sequences in Chapter 9, and Proposition 7.2 is referenced again in 9.1.2. Chapters 9 and 10 also contain many intermediate results that resemble Proposition 7.2. Aside from this main result, the rest of this chapter consists of preliminary observations and is self-contained.

We define the following sequences in terms of a system of nested recurrences:

**Definition 7.1.** *Define sequences $R(n)$ for $n \geq 1$, $S(n)$ for $n \geq 0$, and $T(n)$ for $n \geq 0$:*

- $R(1) = 1$, $R(2) = 2$, $R(n) = R(n - R(n-1)) + S(n-1)$ *for* $n \geq 3$
- $S(0) = 1$, $S(1) = 1$, $S(n) = S(n - R(n)) + S(n - R(n-1))$ *for* $n \geq 2$
- $T(0) = 1$, $T(n) = T(n - R(n)) + T(n - S(n))$ *for* $n \geq 1$

These are sequences A272611, A272612, and A272613 respectively in OEIS [31]. They appear to behave fairly chaotically, and, much like the $Q$-sequence, it is unknown whether they live, persist, or strongly die. Empirically, $R(n)$ grows approximately like $\frac{n}{2}$, $S(n)$ grows like $\frac{n}{4}$ and $T(n)$ grows like $n^{\alpha}$ for some $\alpha$ strictly between 1 and 2. It would make some sense if $\alpha$ were the number such that $4^{\alpha} = 2^{\alpha} + 3^{\alpha}$ (approximately

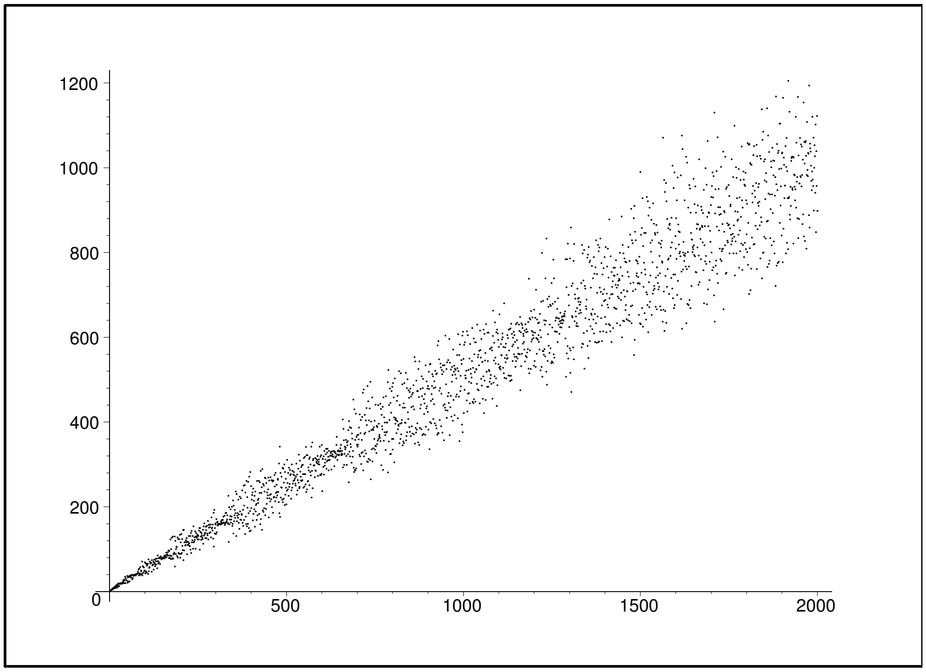


Figure 7.1: Plot of $R(1)$ through $R(2000)$

1.507), though it is unclear if this is actually the case. For plots of these sequences, see Figures 7.1, 7.2, and 7.3 respectively.

The $R$, $S$, and $T$ sequences are of interest because we can generate them with the Hofstadter $Q$-recurrence. We make the following observation:

**Proposition 7.2.** *Let $K \geq 0$, $\lambda \geq 9$ and $\mu \geq K+6$ be integers. The initial condition $\langle a_1, a_2, \ldots, a_K, 5, \lambda, 4, \mu \rangle$ (each $a_i$ an arbitrary integer) for the Hofstadter $Q$-recurrence generates the following pattern for indices $n \geq K+5$,*

- $Q_T(K+5k) = 5R(k)$
- $Q_T(K+5k+1) = 5S(k)$
- $Q_T(K+5k+2) = \lambda T(k)$
- $Q_T(K+5k+3) = 4$
- $Q_T(K+5k+4) = 5R(k)$.

*The pattern lasts as long as the $R$, $S$, and $T$ sequences live and as long as $\lambda T(k) \geq K+5k+4$.*

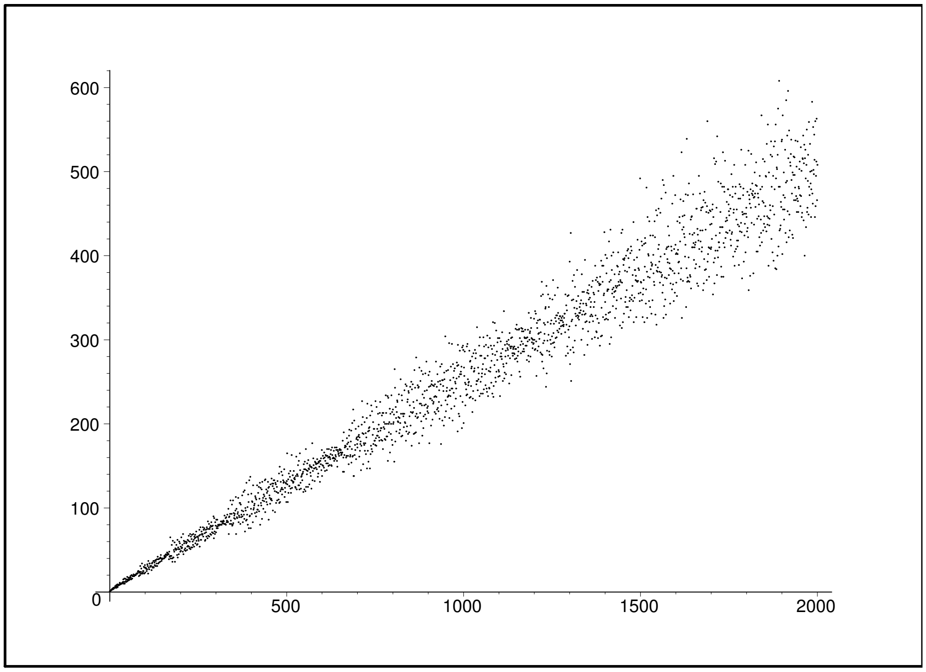


Figure 7.2: Plot of $S(0)$ through $S(2000)$

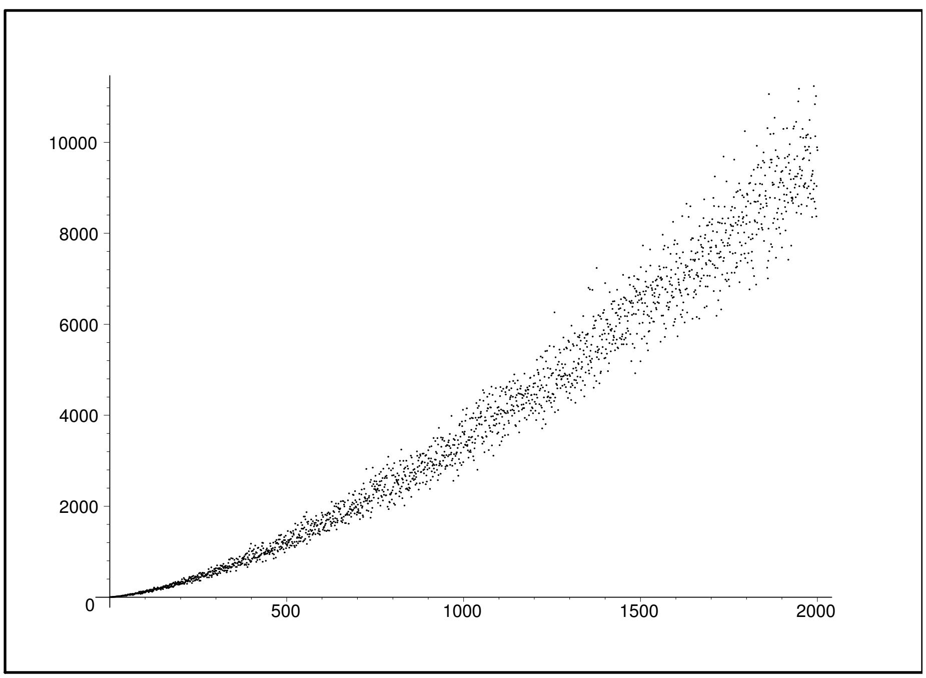


Figure 7.3: Plot of $T(0)$ through $T(2000)$

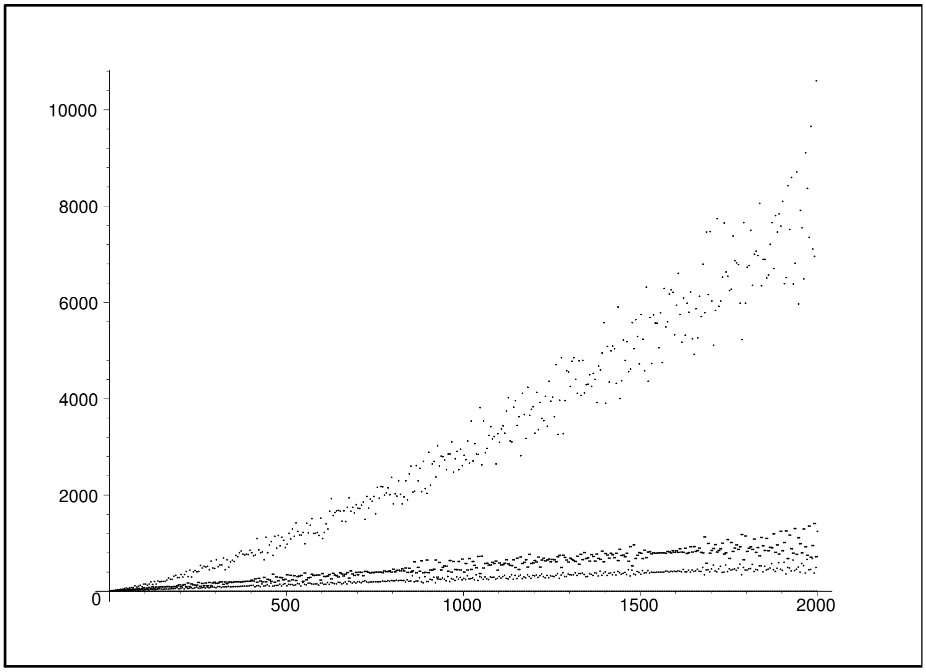


Figure 7.4: The first 2000 terms of A272610 (Initial Condition $\langle 5, 9, 4, 6 \rangle$)

The condition $\lambda T(k) \geq K + 5k + 4$ may seem somewhat restrictive, but it appears to be satisfied for sufficiently large $\lambda$. If $K = 0$, experimental evidence indicates that $\lambda = 9$ suffices. It fails for $\lambda = 8$, as $T(10) = 6$, and it is not true that $48 \geq 54$. The case $K = 0$, $\lambda = 9$ and $\mu = 6$, A272610 in OEIS, is depicted in Figure 7.4.

*Proof.* The proof is by induction on $n$. As a base case, we first manually check $n = K+5$ through $n = K + 8$.

- $Q_T(K+5) = Q_T(K+5-\mu) + Q_T(K+5-4) = Q_T(K+1) = 5 = 5R(1)$.
- $Q_T(K+6) = Q_T(K+6-5) + Q_T(K+6-\mu) = Q_T(K+1) = 5 = 5S(1)$.
- $Q_T(K+7) = Q_T(K+7-5) + Q_T(K+7-5) = 2Q_T(K+2) = 2\lambda = \lambda T(1)$.
- $Q_T(K+8) = Q_T(K+8-2\lambda) + Q_T(K+8-5) = Q_T(K+3) = 4$.

We now proceed by induction on $n$ for $n \geq K + 9$. There are 5 cases to consider.

**$K + n \equiv 0 \pmod{5}$:** Here, $n = K + 5k$ for some $k \geq 2$. We have

$$\begin{aligned}
Q_T(K+5k) &= Q_T(K+5k-Q_T(K+5k-1)) \\
&\quad + Q_T(K+5k-Q_T(K+5k-2)) \\
&= Q_T(K+5k-5R(k-1)) + Q_T(K+5k-4) \\
&= 5R(k-R(k-1)) + 5S(k-1) \\
&= 5R(k),
\end{aligned}$$

as required.

**$K + n \equiv 1 \pmod{5}$:** Here, $n = K + 5k + 1$ for some $k \geq 2$. We have

$$\begin{aligned}
Q_T(K+5k+1) &= Q_T(K+5k+1-Q_T(K+5k)) \\
&\quad + Q_T(K+5k+1-Q_T(K+5k-1)) \\
&= Q_T(K+5k+1-5R(k)) + Q_T(K+5k+1-5R(k-1)) \\
&= 5S(k-R(k)) + 5S(k-R(k-1)) \\
&= 5S(k),
\end{aligned}$$

as required.

**$K + n \equiv 2 \pmod{5}$:** Here, $n = K + 5k + 2$ for some $k \geq 2$. We have

$$\begin{aligned}
Q_T(K+5k+2) &= Q_T(K+5k+2-Q_T(K+5k+1)) \\
&\quad + Q_T(K+5k+2-Q_T(K+5k)) \\
&= Q_T(K+5k+2-5S(k)) + Q_T(K+5k+2-5R(k)) \\
&= \lambda T(k-S(k)) + \lambda T(k-R(k)) \\
&= \lambda T(k),
\end{aligned}$$

as required.

$K+n \equiv 3 \pmod 5$**:** Here, $n = K+5k+3$ for some $k \geq 2$. We have

$$\begin{aligned} Q_T(K+5k+3) &= Q_T(K+5k+3-Q_T(K+5k+2)) \\ &\quad + Q_T(K+5k+3-Q_T(K+5k+1)) \\ &= Q_T(K+5k+3-\lambda T(k)) + Q_T(K+5k+3-5S(k)) \\ &= 0+4 \\ &= 4, \end{aligned}$$

as required.

$K+n \equiv 4 \pmod 5$**:** Here, $n = K+5k+4$ for some $k \geq 1$. We have

$$\begin{aligned} Q_T(K+5k+4) &= Q_T(K+5k+4-Q_T(K+5k+3)) \\ &\quad + Q_T(K+5k+4-Q_T(K+5k+2)) \\ &= Q_T(K+5k+4-4) + Q_T(K+5k+4-\lambda T(k)) \\ &= Q_T(K+5k)+0 \\ &= 5R(k), \end{aligned}$$

as required.

What assumptions do we make about $\lambda$ and $\mu$? When computing $Q_T(K+6)$, we definitely require $\mu \geq 6$. After this, we never see $\mu$ again. As far as $\lambda$ goes, when computing $Q_T(K+5k+3)$, we need $\lambda T(k) \geq K+5k+4$ for every $k$, as required. $\square$

Proposition 7.2 illustrates that we can (at least for awhile, and conjecturally forever) generate an unusual solution to the Hofstadter $Q$-recurrence consisting of five interleaved subsequences. Four of the subsequences appear to behave chaotically, but one of the subsequences is constant (fours, in this case). This begs the question: do other, similar solutions exist for the $Q$-recurrence (and for other recurrences)? The Maple code `http://github.com/nhf216/thesis/RSTsearch.txt` is designed to help search for more of them. The main procedure in this package tries to find a solution to the $Q$-recurrence consisting of $m$ interleaved sequences, one of which is a constant sequence with every term equal to $bm-1$ for some constant $b \geq 1$. It does this by generating the sequence with a certain symbolic initial condition. (See Chapter 8 for an introduction to symbolic initial conditions.) Manual inspection of the resulting solution and bounds on symbols lead to conjectures. In particular, if generating more terms does not cause the sequence to strongly die and does not increase the lower bound required on the symbolic terms, this is evidence in favor of an $(R, S, T)$-like solution. (This is analogous to how, in Proposition 7.2, we obtain bounds on symbols that we are able to satisfy.)

Our conjecture is the following: Exploration yields the following empirical observations for $m \geq 4$:

- If $m \equiv 2 \pmod 3$, then there appears to be a solution to the Hofstadter $Q$-recurrence with $m$ interleaved sequences, including one constant $m-1$ sequence. (Proposition 7.2 gives an example of such a sequence. Another example is given by the initial condition

  $$\langle 7, N, 8, 8, 8, 8, 8, N, 7, 8, 16, 16, 16, 16, 16, N \rangle,$$

  as long as $N \geq 26$. See Figures 7.5 and 7.6.) As seen in Figure 7.6, some of the other interleaved sequences are quite chaotic, others are mildly chaotic, and others (beyond the $m-1$ sequence) are predictable.

- For every integer $b \geq 2$, there appears to be a solution to the Hofstadter $Q$-recurrence with 5 interleaved sequences, each one chaotic except for one constant $5b-1$ sequence. (For example, the initial condition

  $$\langle 9, N, 5, 5, N, 9, N, 5, 5, N, 9, 5, 10, 10, N \rangle$$

  seems to generate one of these with $b = 2$, as long as $N \geq 20$. See Figure 7.7.)

- If $m \equiv 0 \pmod 3$, then for every integer $b \geq 2$, there appears to be a finite, but fairly long-lasting, solution to the Hofstadter $Q$-recurrence with $m$ interleaved sequences, each one chaotic except for one constant $bm-1$ sequence. (For example, the initial condition

  $$\langle 11, N, 6, 6, 6, N, 11, N, 6, 6, 6, N, 11, 6, 12, 12, 12, N \rangle$$

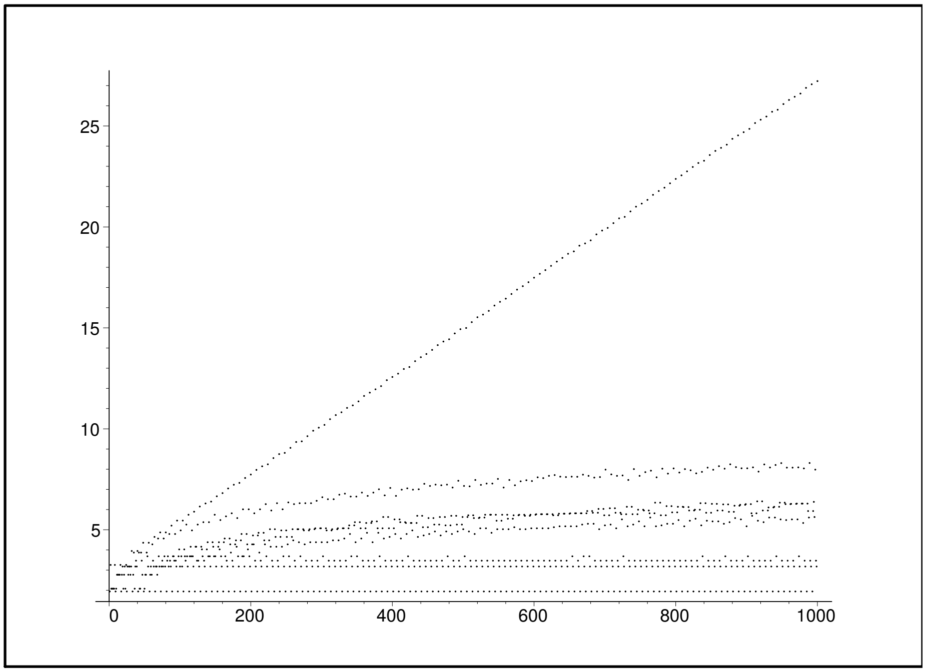


Figure 7.5: First 1000 terms of Hofstadter $Q$-recurrence with initial condition $\langle 7, 26, 8, 8, 8, 8, 8, 26, 7, 8, 16, 16, 16, 16, 16, 26 \rangle$ (A284054), log plot

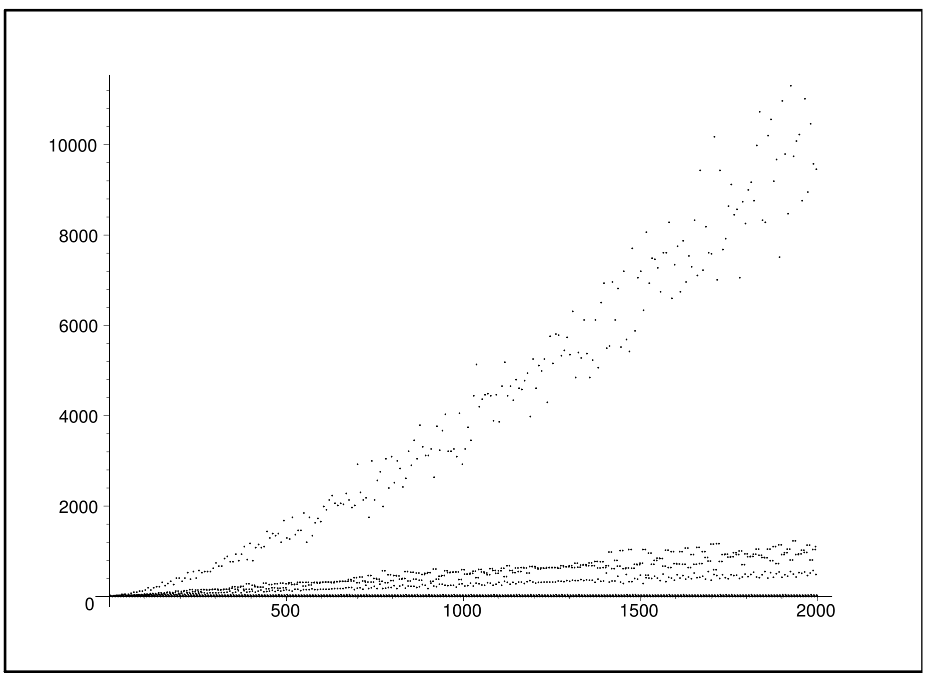


Figure 7.6: First 2000 terms of Hofstadter $Q$-recurrence with initial condition $\langle 7, 26, 8, 8, 8, 8, 8, 26, 7, 8, 16, 16, 16, 16, 16, 26 \rangle$ (A284054), giant terms removed

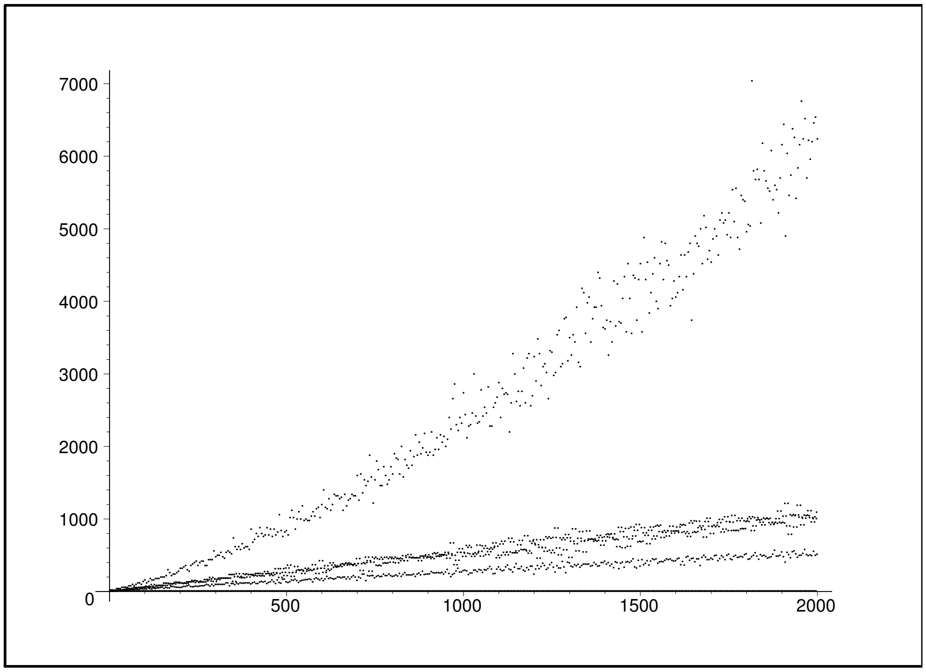


Figure 7.7: First 2000 terms of Hofstadter $Q$-recurrence with initial condition $\langle 9, 20, 5, 5, 20, 9, 20, 5, 5, 20, 9, 5, 10, 10, 20\rangle$ (A284053)

generates one of these of length 2179 with $m = 6$ and $b = 2$, as long as $N \geq 20$. See Figure 7.8.)

- If $m \equiv 1 \pmod 3$, then there appears to be a temporary solution to the Hofstadter $Q$-recurrence with $m$ interleaved sequences, each one chaotic except for one constant $m - 1$ sequence. The duration of this solution depends on how large the symbolic terms in the initial condition are. (For example, the initial condition

$$\langle 6, N, 7, 7, 7, 7, N, 6, 7, 14, 14, 14, 14, N\rangle$$

generates such a solution with $m = 7$. As $N$ is taken larger, the pattern lasts longer.)

- No solution of this sort arises in this exploration for any values of $m$ and $b$ not described above.

Notably, these observations suggest that $m = 5$ is the only case where we can take $b \geq 2$ and obtain infinite sequences. The reason for this and, more generally, the reasons

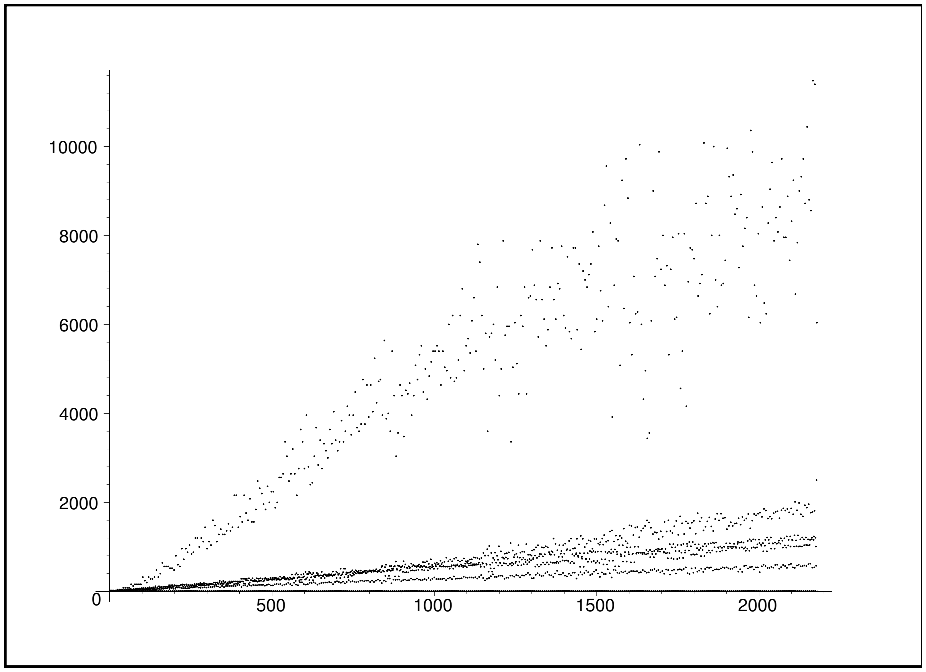


Figure 7.8: All 2179 terms of Hofstadter $Q$-recurrence with initial condition $\langle 11, 20, 6, 6, 6, 20, 11, 20, 6, 6, 6, 20, 11, 6, 12, 12, 12, 20\rangle$ (A283903)

for all of the cases, remain mysterious. Also, no solutions like these are known for any other recurrences, though perhaps they do exist.

We conclude by questioning whether the $R$, $S$, and $T$ sequences are the only chaotic but seemingly stable solutions to their own system of nested recurrences. The initial condition $\langle 4, N, 5, 5, N, 4, 5, 10, 10, N\rangle$ (where $N$ is some large integer) generates a solution similar to $Q_T$ from Proposition 7.2, but the chaotic sequences, at first glance, are slightly different. But, further observation shows that the original $R$ and $S$ sequences do appear, but they start later (around index 28). The corresponding $T$-like sequence is different, though.

# Chapter 8

# Background on Parametrized Families of Initial Conditions

The work of Chapters 3, 4, 5, 6, and, to some extent, 7, as well as most of the literature (e.g. [15,30,32]), is primarily concerned with finding solutions to nested recurrences where the solutions have some specific properties. Given a desired behavior, the goal has been to find an initial condition that realizes that behavior. In Chapter 9 and in Chapter 10, we flip the conventional process on its head, and this chapter gives the background we need for the analysis in those chapters. Instead of focusing on a particular style of solution, we now focus on a particular style of initial condition. From there, we try to characterize the resulting behaviors. Usually, the solutions die (either weakly or strongly; we have results for both notions). The theorems tend to be of the flavor, "For all sufficiently long/large initial conditions of a given type, the sequence defined by some recurrence weakly/strongly dies in a specific way."

We would like to be able to consider infinite families of similar initial conditions simultaneously. So, we consider initial conditions that contain one or more symbolic parameters. Each setting of the parameters then gives an initial condition. For example, Chapter 9 is an extensive study of the family of initial conditions $\langle 1, 2, 3, \ldots, N\rangle$, where $N$ is a parameter. If we specialize $N$ to be, say, $N = 7$, the initial condition becomes $\langle 1, 2, 3, 4, 5, 6, 7\rangle$. Generally speaking, our results hold when the parameters are sufficiently large, and each result includes a description of what "sufficiently large" means.

Going forward, there are four different approaches to viewing the parameters in our initial conditions.

**Large Integers:** As we just mentioned, our results apply when the parameters are sufficiently large integers. Consequently, it is most straightforward to treat the parameters as large integers whose values have yet to be determined. The primary disadvantage of this point of view is that, a priori, we do not know what "sufficiently large" means; our process of discovering and proving theorems, as a consequence, defines this term in each case. The other viewpoints do not suffer from this shortcoming, though they are less natural.

**Symbols:** When doing computations involving unknown parameters, our procedures treat them as symbols. But, plain symbols do not contain enough information to compute everything. So, we must endow our symbols with the following auxiliary information:

- A symbol $N$ is treated as larger than every natural number. (This successfully circumnavigates the "sufficiently large" concern.)
- A symbol $N$ may be assigned to congruence classes modulo some positive integers. (These congruence classes must be consistent with one another. For example, we cannot simultaneously have $N \equiv 1 \pmod 2$ and $N \equiv 6 \pmod 8$.) This is necessary because computing terms of sequences with parameters in them frequently depends on arithmetic properties of $N$. (See any theorem on strong death in Chapter 9 for an example.)
- Two different symbols $N_1$ and $N_2$ appearing in the same initial condition are given a relative ordering, say, $N_1 < N_2$. Furthermore, $N_2$ is treated as larger than any expression built only out of $N_1$'s (and real numbers).

**Nonstandard Integers:** The properties described above that we need to assign to our symbols are reminiscent of nonstandard arithmetic [25]. The existence of nonstandard models of arithmetic is implied by Gödel's Incompleteness Theorems [14]. Such models must contain "infinite" numbers, that is, natural numbers that are larger than any standard natural number. As natural numbers, these infinite numbers are endowed with intrinsic arithmetic properties. Most notably, for any natural numbers $N$ and $M$ (standard or nonstandard), $N \bmod M$ is defined. (In particular, any two congruence classes that $N$ falls into must be mutually consistent.) The theorems we prove in the next two chapters can be thought of as theorems about nonstandard integer sequences (where indices and terms can each be nonstandard).

**$p$-Adic Integers:** For an integer $p \geq 2$, a $p$-adic integer [17] is a formal sum $\sum_{k=0}^{\infty} a_k p^k$ for some integers $a_k$, each ranging from 0 to $p-1$ (inclusive). A generic $p$-adic integer is often written as a left-nonterminating string of digits: $\cdots a_4 a_3 a_2 a_1 a_0{}_p$. The subscript $p$ means that it is $p$-adic; if $p$ is clear from context, the subscript is often omitted. The natural numbers are contained in the $p$-adic integers. If the sequence $(a_k)_{k\geq 0}$ is eventually zero, then the resulting $p$-adic integer is a natural number (written in its base-$p$ representation). But there are uncountably many $p$-adic integers. Addition and multiplication of $p$-adic integers is done as if they were base-$p$ integers.

Negative integers are also $p$-adic integers, as we can write

$$-1 = \cdots (p-1)(p-1)(p-1)_p .$$

(One way to see this is to formally sum this as a geometric series; another way is to add 1 and observe how the carried digits cause the result to become zero.) In this way, the $p$-adic integers form a ring, and, if $p$ is prime, they form an integral domain. Some fractions are also $p$-adic integers. (For example, $-\frac{1}{2} = \cdots 111111_3$.)

At times, it can be helpful to think of our initial condition parameters as $p$-adic integers. In our main result of Section 9.1.2, the eventual behavior of a sequence depends on the value of a parameter $N$ mod higher and higher powers of 5. If $N$ is a 5-adic integer, it is easy to work with these congruences, as they are obtained from the low-order digits of the 5-adic expansion of $N$.

A disadvantage of working with $p$-adic integers is the fact that fractions and negative numbers can be $p$-adic integers. Important to our analysis is that the natural numbers are totally ordered and that there is no largest natural number. But, the $p$-adic integers violate these properties as a whole. That being said, we still sometimes refer to parameters as $p$-adic integers when it appears that they do not end up corresponding to negative integers and when we only mod by powers of $p$. (The values of our sequences at negative integers is given on page 7 in the definition of strong death, and taking a $p$-adic integer mod something other than a power of $p$ is a messy operation.)

In Chapter 9, we consider various recurrences under initial conditions of the form $\langle 1, 2, 3, \ldots, N\rangle$. In Chapter 10, we consider initial conditions of the form $\langle N, 2\rangle$, $\langle 2, N\rangle$, $\langle N, 4, N, 4\rangle$, and $\langle 4, N, 4, N\rangle$. In Section 8.1, we discuss in more detail the general method we use in these next two chapters.

## 8.1 The General Method

We now describe the general method of analyzing sequences generated by a given parametrized family of initial conditions for a nested recurrence. The same procedure works for either notion of death, so when we say a sequence dies, we mean it dies according to whichever convention we are choosing to work with. This process is implemented in the Maple package `http://github.com/nhf216/thesis/nonstdhof.txt` in the procedures `ProveEventualSolution` and `ProveLongTermEventualSolution`.

1. At the start, we are given a recurrence and a parametrized initial condition family. The goal is to determine the values of all terms in the sequence. If, when doing so, we realize that the sequence must refer to an illegal index at some point, then we additionally conclude that the sequence dies. If the recurrence under consideration is basic, then strong death is synonymous with referencing a zero or negative term when computing a sequence value.

2. Generate the next terms of the sequence. It may be necessary to specify certain properties of the parameters in order to generate these terms. (These can either be specified at the outset, or we can iterate through all possibilities.) Keep generating terms until one of the following things happens:

    - The sequence dies.

- A regular pattern (a la Golomb [15] or Ruskey [30]) emerges.

3. If the sequence dies, then we are done. Otherwise, try to prove by induction that the observed pattern actually exists for some amount of time. If this succeeds, carefully examine the assumptions made in the proof and try to extract from it the first index where the pattern no longer persists. (This will typically be some function of the parameters, though it will sometimes be $\infty$, as patterns can persist forever.) On the other hand, if the proof of the pattern's persistence fails, return to step 2 and generate more terms.

4. Return to step 2, replacing the previously-considered initial condition by all of the sequence values that we now know about.

In a sense, this process is a close relative of the main algorithm in Chapter 3, as it can be used to show that a solution to a nested recurrence is eventually an interleaving of linear-recurrent sequences. But, this process requires a specific family of initial conditions as an input, and it does not discover infinite families of solutions.

This process is fairly simple to describe, and there are many applications of it in the upcoming chapters. But, the implementation requires navigating many messy details. The difficulties arise from the following sources:

- We need to keep track of arithmetic properties of symbols.
- We need to compare symbolic expressions to see which is larger.
- More generally, given a partition of the (nonstandard) integers into intervals (each corresponding to an interval where a certain pattern exists), we need to be able to determine which interval a given expression falls into.
- We need to automatically look for patterns, which requires a precise notion of what a "pattern" is.
- We need to be able to algebraically determine where patterns come to an end.

None of these tasks is particularly straightforward, and, as a result, the implementation is fairly involved.

Sometimes, the sequences we discover consist of "patterns of patterns," meaning that applying this method to such a sequence will never tell the whole story. Rather, the patterns we see in the sequence keep recurring in some form. In order to properly analyze such solutions, it would be necessary to employ a second-order version of this methodology that searches for such meta-patterns. (For an example of a sequence family with a meta-pattern, see p. 150.) Of course, why stop at second order? This sort of thing could continue for arbitrarily many orders. But, our implementation currently only natively handles first order patterns, though, as is done in Chapter 10, the procedure can sometimes be adapted to handle specific cases of higher-order patterns.

# Chapter 9

# Nested Recurrences with Initial Conditions 1 through $N$

The Hofstadter $Q$-sequence appears to be approximately linear. Keeping this in mind, it may be fruitful to examine the $Q$-recurrence under linear initial conditions. Since our initial conditions must contain only integers, an obvious first initial condition to try is $\langle 1, 2, 3, \ldots, N \rangle$. In Section 9.1, we use the process described in Chapter 8 to explore the sequences generated by the $Q$-recurrence under these initial conditions. In Sections 9.2 and 9.3, we explore related recurrences under the same initial conditions. Finally, in Section 9.4, we use these initial conditions and the methodology of Chapter 8 to prove that certain sequences that naturally generalize the slow sequence in Chapter 6 are *not* slow.

## 9.1 The Hofstadter $Q$-Recurrence

In this section, we consider sequences obtained from the Hofstadter $Q$-recurrence and an initial condition of the form $\langle 1, 2, 3, \ldots, N \rangle$, the result of which we denote by $Q_N$. Observe that $Q_2$ is precisely the $Q$-sequence with all terms shifted left by one index. (This follows almost immediately from Proposition 2.7.)

### 9.1.1 Weak Death

Our first result is a description of the behavior of the sequences $Q_N$ under weak death.

**Theorem 9.1.** *For $N = 8$, $N = 11$, $N = 12$, or $N \geq 14$, $Q_N$ weakly dies.*

The proof of Theorem 9.1 will serve as a basic illustration of the first part of our general method from Chapter 8.

*Proof.* It is straightforward to verify that $Q_8(420) = 430$, $Q_{11}(199) = 206$, and $Q_{12}(69) = 77$, so these sequences all weakly die [31, A278060,A278063,A278064].

Now, suppose $N \geq 14$. We compute 28 terms following the initial condition and obtain values

$$3, N+1,\ N+2,\ 5,\ N+3,\ 6,\ 7,\ N+4,\ N+6,\ 10,\ 8,\ N+6,\ N+10,\ 12,\ N+7,\ 14,$$
$$N+12,\ 11,\ N+11,\ N+15,\ 16,\ 13,\ 17,\ 15,\ N+14,\ 20,\ 20,\ 2N+8$$

See Appendix B for explicit computations of these terms.

Note that, as we compute these values, we use the fact that $Q_N(i) = i$ for $i \leq 13$. But, $Q_N(13)$ (when computing $Q_N(N+27)$) is the largest term we evaluate whose argument does not include an $N$. Now, observe that $2N+8 \geq N+1$ whenever $N \geq 21$. This means that, if $N \geq 21$, then $Q_N(N+29)$ will fail to exist (since we are considering weak death). So, the sequence weakly dies whenever $N \geq 21$.

This just leaves the values $14 \leq N \leq 20$ to examine. This is a finite range, so it suffices to check all of these values individually. But, these seven sequences all weakly die according to the same pattern, and our proof below will illustrate this. Suppose $14 \leq N \leq 20$. We now compute four more terms.

$$\begin{aligned}
\mathbf{Q_N(N+29)} &= Q_N(N+29-Q_N(N+28)) + Q_N(N+29-Q_N(N+27)) \\
&= Q_N(N+29-(2N+8)) + Q_N(N+29-(20)) \\
&= Q_N\left(-N+21\right) + Q_N\left(N+9\right) = -N+21+N+6 = \mathbf{27}
\end{aligned}$$

$$\begin{aligned}
\mathbf{Q_N(N+30)} &= Q_N(N+30-Q_N(N+29)) + Q_N(N+30-Q_N(N+28)) \\
&= Q_N(N+30-(27)) + Q_N(N+30-(2N+8)) \\
&= Q_N\left(N+3\right) + Q_N\left(-N+22\right) = N+2-N+22 = \mathbf{24}
\end{aligned}$$

$$\begin{aligned}
\mathbf{Q_N(N+31)} &= Q_N(N+31-Q_N(N+30)) + Q_N(N+31-Q_N(N+29)) \\
&= Q_N(N+31-(24)) + Q_N(N+31-(27)) \\
&= Q_N\left(N+7\right) + Q_N\left(N+4\right) = 7+5 = \mathbf{12}
\end{aligned}$$

$$
\begin{aligned}
\mathbf{Q_N(N+32)} &= Q_N(N+32-Q_N(N+31)) + Q_N(N+32-Q_N(N+30)) \\
&= Q_N(N+32-(12)) + Q_N(N+32-(24)) \\
&= Q_N\left(N+20\right) + Q_N\left(N+8\right) = N+15+N+4 = \mathbf{2N+19}.
\end{aligned}
$$

Here we use the facts that $21-N$ and $22-N$ both lie in the initial condition. We now observe that $2N+19 \geq N+33$ whenever $N \geq 14$. This means that, if $14 \leq N \leq 20$, then $Q_N(N+33)$ fails to exist. So, the sequence weakly dies whenever $14 \leq N \leq 20$, as required. $\square$

Theorem 9.1 says that $Q_N$ weakly dies for all but finitely many $N$. This begs the question of what happens when $N \in \{2,3,4,5,6,7,9,10,13\}$. We have already mentioned that the case $N = 2$ is Hofstadter's sequence shifted by 1 position, so it is unknown whether $Q_2$ dies (see Chapter 2). Since $Q_2(3) = 3$, $Q_3 = Q_2$, and we have that $N = 3$ also gives Hofstadter's sequence. The remaining $N$ values in this set give sequences that are different from Hofstadter's sequence and different from each other. Like Hofstadter's, it is unknown whether any of these sequences dies. All of these sequences last for at least 30 million terms. See Appendix C for plots of the first 2000 terms of all eight sequences. Observe that, as $N$ increases, the plots appear to progress from the characteristic "sausage" pattern of the Hofstadter $Q$-sequence to more of a wedge shape.

### 9.1.2 Strong Death

#### General Structure of $Q_N$

We now examine what happens to the sequences $Q_N$ under the strong death convention. Surprisingly, the behavior beyond the weak death point depends on the congruence class of $N$ modulo 5. For fixed $N$, we define the following sequences:

**Definition 9.2.** *Define* $A_0 = N-2$, $A_1 = 2N+4$ *and*

$$B_1 = -11N - 22.$$

*Then, for $i \geq 1$, define*

$$A_{i+1} = A_i\left(\frac{A_i - A_{i-1} + 2}{5}\right) + B_i$$

*and $B_{i+1} = A_{i+1} - A_i$. Next, define $C_1 = (N-1) \bmod 5$, and for $i \geq 2$, define $C_i = (A_i + 2i + 1) \bmod 5$. Finally, for all $i \geq 1$, define $C_i' = \max(0, ((3 - C_i) \bmod 5) - 1)$.*

Also, recall the $R$, $S$, and $T$ sequences from Chapter 7. We have the following theorem:

**Theorem 9.3.** *Let $N$ be a natural number (standard, nonstandard, or nonnegative and 5-adic). Let $j$ be the first index where $C_j \neq 1$ (or $j = \infty$ if $C_j = 1$ for all $j$). Provided $N \geq 35$, the sequence $Q_N(n)$ has the following properties:*

- *For all $1 \leq i \leq N$, $Q_N(i) = i$.*
- *For $1 \leq k \leq 28$, $Q_N(N+k)$ is as in Theorem 9.1. The next six terms are $Q_N(N+29) = N+6$, $Q_N(N+30) = 24$, $Q_N(N+31) = 32$, $Q_N(N+32) = 2N+4$, $Q_N(N+33) = 3$, $Q_N(N+34) = 32$. Thereafter, for $35 \leq 5k + r \leq A_1 + C_1'$ with $0 \leq r < 5$,*
  - $Q_N(N+5k) = A_1 k + B_1$
  - $Q_N(N+5k+1) = 5$
  - $Q_N(N+5k+2) = A_1$
  - $Q_N(N+5k+3) = 3$
  - $Q_N(N+5k+4) = 5$
- *For each $1 \leq m < j$, $Q_N(A_m + 2) = 5$, $Q_N(A_m + 3) = 8$, $Q_N(A_m + 4) = A_{m+1}$, $Q_N(A_m + 5) = 3$, $Q_N(A_m + 6) = 8$, and for all $7 \leq 5k + r \leq A_{m+1} + C_{m+1}'$ with $0 \leq r < 5$,*
  - $Q_N(A_m + 5k) = 3$
  - $Q_N(A_m + 5k + 1) = 5$
  - $Q_N(A_m + 5k + 2) = A_{m+1} k + B_{m+1}$

- $Q_N(A_m + 5k + 3) = 5$
- $Q_N(A_m + 5k + 4) = A_{m+1}$

- *If* $C_j = 0$ *and* $N \geq 118$*, then* $Q_N$ *strongly dies after* $A_j + 160$ *terms. See Appendix D for the remaining* $158$ *terms.*

- *If* $C_j = 2$*, then* $Q_N(A_j + 1) = 4$*,* $Q_N(A_j + 2) = A_j\left(\frac{A_j - A_{j-1} - 4}{5}\right) + B_j + 2$*, and thereafter, for* $5k + r \geq 3$ *with* $0 \leq r < 5$

  - $Q_N(A_j + 5k) = A_j T(k)$
  - $Q_N(A_j + 5k + 1) = 4$
  - $Q_N(A_j + 5k + 2) = 5R(k)$
  - $Q_N(A_j + 5k + 3) = 5R(k + 1)$
  - $Q_N(A_j + 5k + 4) = 5S(k + 1)$

  *assuming the* $R$*,* $S$*, and* $T$ *sequences from Chapter 7 do not weakly die and assuming the* $T$ *sequence stays large enough.*

- *If* $C_j = 3$*, then* $Q_N$ *strongly dies after* $A_j + 4$ *terms. The remaining* $4$ *terms are:*

  - $Q_N(A_j + 1) = 6$
  - $Q_N(A_j + 2) = A_j + 5$
  - $Q_N(A_j + 3) = A_j\left(\frac{A_j - A_{j-1} - 5}{5}\right) + B_j$
  - $Q_N(A_j + 4) = 0$

- *If* $C_j = 4$*, then* $Q_N$ *strongly dies after* $A_j + 14$ *terms. The remaining* $11$ *terms are:*

  - $Q_N(A_j + 4) = 7$
  - $Q_N(A_j + 5) = A_j + 5$
  - $Q_N(A_j + 6) = 4$
  - $Q_N(A_j + 7) = A_j + 2$
  - $Q_N(A_j + 8) = 13$

- $Q_N(A_j+9) = A_j\left(\frac{A_j - A_{j-1} - 6}{5}\right) + B_j + 7$
- $Q_N(A_j+10) = 5$
- $Q_N(A_j+11) = 4$
- $Q_N(A_j+12) = A_j + 15$
- $Q_N(A_j+13) = A_j\left(\frac{A_j - A_{j-1} - 6}{5}\right) + B_j + 7$
- $Q_N(A_j+14) = 0$

The proof of Theorem 9.3 requires the following lemma:

**Lemma 9.4.** *Let $K \geq 0$ be an integer, and let $\lambda$ and $\mu$ be any integers satisfying $\lambda > K+5$ and $\lambda + \mu > K+6$. Then, for arbitrary integers $a_1, a_2, \ldots, a_K$, the initial condition $\langle a_1, a_2, \ldots, a_K, \mu, 5, \lambda, 3\rangle$ generates the pattern*

- $Q_C(K+5k) = 5$
- $Q_C(K+5k+1) = \lambda k + \mu$
- $Q_C(K+5k+2) = 5$
- $Q_C(K+5k+3) = \lambda$
- $Q_C(K+5k+4) = 3$

*satisfies the Hofstadter Q-recurrence from $Q_C(K+1)$ through $Q_C(\lambda)$.*

*In addition, the pattern can only end when computing the 3 or the second 5 (the terms referring to $\lambda$), so the pattern may extend to index $Q_C(\lambda+i)$ for some $i \leq 3$ (depending on the congruence class of $\lambda$ mod 5).*

*Proof.* The proof is by induction on the index. The base cases are $Q_C(K+1)$ through $Q_C(K+4)$, which are part of the initial condition. Now, suppose $K+5 \leq n \leq \lambda$, and suppose that $Q_C(n')$ is what we want for all $K+1 \leq n' < n$. There are five cases to consider:

$n - K \equiv 0 \pmod 5$**:** In this case, $n = K + 5k$ for some $k$. Applying the $Q$-recurrence, we have

$$\begin{aligned} Q_C(K+5k) &= Q_C(K+5k - Q_C(K+5k+4)) \\ &\quad + Q_C(K+5k - Q_C(K+5k+3)) \\ &= Q_C(K+5k-3) + Q_C(K+5k-\lambda) \\ &= 5 + 0 \\ &= 5, \end{aligned}$$

as required.

$n - K \equiv 1 \pmod 5$**:** In this case, $n = 5k + 1$ for some $k$. Applying the $Q$-recurrence, we have

$$\begin{aligned} Q_C(K+5k+1) &= Q_C(K+5k+1 - Q_C(K+5k)) \\ &\quad + Q_C(K+5k+1 - Q_C(K+5k-1)) \\ &= Q_C(K+5k+1-5) + Q_C(K+5k+1-3) \\ &= \lambda(k-1) + \mu + \lambda \\ &= \lambda k + \mu, \end{aligned}$$

as required.

$n - K \equiv 2 \pmod 5$**:** In this case, $n = K+5k+2$ for some $k$. Applying the $Q$-recurrence, we have

$$\begin{aligned} Q_C(K+5k+2) &= Q_C(K+5k+2 - Q_C(K+5k+1)) \\ &\quad + Q_C(K+5k+2 - Q_C(K+5k)) \\ &= Q_C(K+5k+2-(\lambda k+\mu)) + Q_C(K+5k+2-5) \\ &= 0 + 5 \\ &= 5, \end{aligned}$$

as required.

$n - K \equiv 3 \pmod 5$**:** In this case, $n = 5k + 3$ for some $k$. Applying the $Q$-recurrence, we have

$$\begin{aligned}
Q_C(K + 5k + 3) &= Q_C(K + 5k + 3 - Q_C(K + 5k + 2)) \\
&\quad + Q_C(K + 5k + 3 - Q_C(K + 5k + 1)) \\
&= Q_C(K + 5k + 3 - 5) + Q_C(K + 5k + 3 - (\lambda k + \mu)) \\
&= \lambda + 0 \\
&= \lambda,
\end{aligned}$$

as required.

$n - K \equiv 4 \pmod 5$**:** In this case, $n = 5k + 4$ for some $k$. Applying the $Q$-recurrence, we have

$$\begin{aligned}
Q_C(5k + 4) &= Q_C(K + 5k + 4 - Q_C(K + 5k + 3)) \\
&\quad + Q_C(K + 5k + 4 - Q_C(K + 5k + 2)) \\
&= Q_C(K + 5k + 4 - \lambda) + Q_C(K + 5k + 4 - 5) \\
&= 0 + 3 \\
&= 3,
\end{aligned}$$

as required.

This completes the proof of the pattern. Examining which terms refer to which other terms gives the extendability observation, as the first and last cases are the only ones that care about the specific value of $\lambda$. $\square$

We now prove Theorem 9.3.

*Proof.* We now refer the reader back to Theorem 9.1 for terms $Q_N(1)$ through $Q_N(N + 28)$. From there, it is easy to compute $Q_N(N + 29)$ through $Q_N(N + 34)$, and each one equals its purported value. We now compute the next five terms:

- $Q_N(N + 35) = Q_N(N + 3) + Q_N(N + 32) = (N + 2) + (2N + 4) = 3N + 6.$
- $Q_N(N + 36) = Q_N(N + 4) = 5.$

- $Q_N(N+37) = Q_N(N+32) = 2N+4 = A_1$.
- $Q_N(N+38) = Q_N(N+33) = 3$.

(Note that $N \geq 35$ is required when computing $Q_N(N+39)$, as we need $2N+4 \geq N+39$. This is the strongest requirement we impose anywhere on the size of $N$.) By Lemma 9.4, taking $K = N+34$, these four terms spawn a period-5 pattern:

- $Q_N(N+34+5k) = 5$
- $Q_N(N+34+5k+1) = (2N+4)\,k + (3N+6)$
- $Q_N(N+34+5k+2) = 5$
- $Q_N(N+34+5k+3) = 2N+4$
- $Q_N(N+34+5k+4) = 3$,

and this pattern persists through at least $Q_N(A_1)$. Shifting indices and recalling the definitions of $A_1$ and $B_1$ allows us to rewrite this pattern as

- $Q_N(N+5k) = A_1 k - B_1$
- $Q_N(N+5k+1) = 5$
- $Q_N(N+5k+2) = A_1$
- $Q_N(N+5k+3) = 3$
- $Q_N(N+5k+4) = 5$,

which is the required form.

In order to complete the proof of the first part of the theorem, we must show that the pattern continues through index $A_1 + C_1'$. The value of $C_1'$ is determined completely by $N \bmod 5$; it is 3 if $N \equiv 0 \pmod 5$, 2 if $N \equiv 1 \pmod 5$, 1 if $N \equiv 2 \pmod 5$, and 0 otherwise. We now determine when the pattern ends in each case.

**$N \equiv 0 \pmod 5$:** Here, $A_1 \equiv 4 \pmod 5$. This means that $Q_N(A_1+1)$ falls into the $Q_N(N+5k)$ case. By Lemma 9.4, the pattern can end only once we reach the

$Q_N(N+5k+3)$ case. So, the pattern persists through $Q_N(A_1+3) = Q_N(A_1+C_1')$, as required.

$N \equiv 1 \pmod 5$**:** Here, $A_1 \equiv 1 \pmod 5$. This means that $Q_N(A_1 + 1)$ falls into the $Q_N(N+5k+1)$ case (since $N$ itself is congruent to 1 mod 5). By Lemma 9.4, the pattern can end only once we reach the $Q_N(5k+3)$ case. So, the pattern persists through $Q_N(A_1+2) = Q_N(A_1+C_1')$, as required.

$N \equiv 2 \pmod 5$**:** Here, $A_1 \equiv 3 \pmod 5$. This means that $Q_N(A_1 + 1)$ falls into the $Q_N(N+5k+2)$ case (since $N$ itself is congruent to 2 mod 5). By Lemma 9.4, the pattern can end only once we reach the $Q_N(5k+3)$ case. So, the pattern persists through $Q_N(A_1+1) = Q_N(A_1+C_1')$, as required.

$N \equiv 3 \pmod 5$**:** Here, $A_1 \equiv 0 \pmod 5$. This means that $Q_N(A_1 + 1)$ falls into the $Q_N(N+5k+3)$ case (since $N$ itself is congruent to 3 mod 5). By Lemma 9.4, the pattern ends immediately. So, the pattern persists through $Q_N(A_1) = Q_N(A_1 + C_1')$, as required.

$N \equiv 4 \pmod 5$**:** Here, $A_1 \equiv 2 \pmod 5$. This means that $Q_N(A_1 + 1)$ falls into the $Q_N(N+5k+4)$ case (since $N$ itself is congruent to 4 mod 5). By Lemma 9.4, the pattern ends immediately. So, the pattern persists through $Q_N(A_1) = Q_N(A_1 + C_1')$, as required.

We now prove the portion of Theorem 9.3 that refers to a parameter $1 \leq m < j$. Suppose inductively that we are considering the value $m < j$, and the theorem is true for $m-1$. In other words, we are at the conclusion of the period-5 pattern that ends at $Q_N(A_m + C_m')$. Since $m < j$, it must be the case that $C_m' = 1$ (as $C_m = 1$ implies $C_m' = 1$). So we must start our examination with $Q_N(A_m + 2)$. Also, we must have $Q_N(A_m + 1) = A_m$. (We include this as part of our inductive hypothesis.)

Technically, $m = 1$ should be treated as a base case. But, we see that $m = 0$ corresponds to the pattern that ends with $Q_N(A_1 + 1) = A_1$. So, we do not actually need to treat $m = 1$ any differently from other $m$ values, and $m = 0$ can serve as our (already proved) base case.

We compute the next 9 terms:

- $Q_N(A_m + 2) = Q_N(2) + Q_N(A_m - 3) = 2 + 3 = 5$.
- $Q_N(A_m + 3) = Q_N(A_m - 2) + Q_N(3) = 5 + 3 = 8$.
- $Q_N(A_m + 4) = Q_N(A_m - 4) + Q_N(A_m - 1)$. We have that $Q_N(A_m - 4) = A_m$. But, $Q_N(A_m - 1) = A_m k + B_m$, where $k = \frac{A_m - A_{m-1} - 3}{5}$. So,

$$\begin{aligned} Q_N(A_m + 4) &= A_m \left(1 + \frac{A_m - A_{m-1} - 3}{5}\right) + B_m \\ &= A_m \left(\frac{A_m - A_{m-1} + 2}{5}\right) + B_m \\ &= A_{m+1}. \end{aligned}$$

  This term is much larger than $A_m$.
- $Q_N(A_m + 5) = Q_N(A_m - 3) = 3$.
- $Q_N(A_m + 6) = Q_N(A_m + 1) = 8$.
- $Q_N(A_m + 7) = Q_N(A_m - 1) + Q_N(A_m + 4)$. We have from before $Q_N(A_m - 1) = A_m k + B_m$, where $k = \frac{A_m - A_{m-1} - 3}{5}$. But, our calculations in the $Q_N(A_m + 4)$ step allow us to write $Q_N(A_m - 1) = A_{m+1} - A_m$. So, $Q_N(A_m + 7) = A_{m+1} - A_m + A_{m+1} = 2A_{m+1} - A_m$.
- $Q_N(A_m + 8) = Q_N(A_m) = 5$.
- $Q_N(A_m + 9) = Q_N(A_m + 4) = A_{m+1}$.
- $Q_N(A_m + 10) = Q_N(A_m + 5) = 3$.

The first five of these terms are what we want. And, by Lemma 9.4, the last four terms generate a period-5 pattern as in the lemma statement (with $K = A_m + 6$). The resulting pattern can be written as

- $Q_N(A_m + 5k) = 3$
- $Q_N(A_m + 5k + 1) = 5$

- $Q_N(A_m + 5k + 2) = A_{m+1}k + A_{m+1} - A_m = A_{m+1}k + B_{m+1}$
- $Q_N(A_m + 5k + 3) = 5$
- $Q_N(A_m + 5k + 4) = A_{m+1}$,

and this pattern persists at least through $Q_N(A_{m+1})$, as required.

We must now show that the pattern continues through index $A_{m_1} + C'_{m+1}$. The argument is similar to the earlier argument we used regarding $A_1 + C'_1$; in short, it suffices to determine which case $Q_N(A_{m+1})$ falls into. This, in turn, requires determining $(A_{m+1} - A_m) \bmod 5$. We know that $C_{m+1} \equiv (A_{m+1} + 2m + 3) \bmod 5$. This means that $A_{m+1} \equiv (C_{m+1} - 2m - 3) \bmod 5$. Similarly, $A_m \equiv (C_m - 2m - 1) \bmod 5$. But, we know that $C_m = 1$. So, $A_m \equiv -2m \bmod 5$. Combining these yields $A_{m+1} - A_m \equiv (C_{m+1} - 3) \bmod 5$. We now examine the different cases.

**$C_{m+1} = 0$:** Here, $Q_N(A_{m+1}) = A_{m+1}\left(\frac{A_{m+1}-A_m-2}{5}\right) + B_{m+1}$, so the pattern persists through $Q_N(A_{m+1} + 2)$. In this case, we have $C'_{m+1} = 2$, as required.

**$C_{m+1} = 1$:** Here, $Q_N(A_{m+1}) = 5$ (the second instance of 5), so the pattern persists through $Q_N(A_{m+1} + 1)$. In this case, we have $C'_{m+1} = 1$, as required.

**$C_{m+1} = 2$:** Here, $Q_N(A_{m+1}) = A_{m+1}$, so the pattern ceases immediately. In this case, we have $C'_{m+1} = 0$, as required.

**$C_{m+1} = 3$:** Here, $Q_N(A_{m+1}) = 3$, so the pattern ceases immediately. In this case, we have $C'_{m+1} = 0$, as required.

**$C_{m+1} = 4$:** Here, $Q_N(A_{m+1}) = 5$ (the first instance of 5), so the pattern persists through $Q_N(A_{m+1} + 3)$. In this case, we have $C'_{m+1} = 3$, as required.

All that remains is to determine the eventual behaviors for $C_j \in \{0, 2, 3, 4\}$.

**$C_j = 0$:** The first unknown term here is $Q_N(A_j + 3)$. We compute the next 158 terms (see Appendix D), and we observe that the sequence strongly dies because $Q_N(A_j + 160) = 0$. Computation of these terms assumes that $N \geq 118$, because computing $Q_N(A_j + 157)$ refers to $Q_N(118)$, which we assume equals 118 (and this is the strongest requirement we use anywhere in the calculations).

$C_j = 2$**:** The first unknown term here is $Q_N(A_j+1)$. We compute the next 2 terms (keeping in mind that $Q_N(A_j) = A_j$ and $Q_N(A_j - 1) = 5$):

- $Q_N(A_j+1) = Q_N(1) + Q_N(A_j - 4) = 1 + 3 = 4$.
- $Q_N(A_j+2) = Q_N(A_j - 2) + Q_N(2) = A_j\left(\frac{A_j - A_{j-1} - 4}{5}\right) + B_j + 2$.

By Proposition 7.2 (and Proposition 2.7), this results in the pattern

- $Q_N(A_j + 5k) = A_j T(k)$
- $Q_N(A_j + 5k + 1) = 4$
- $Q_N(A_j + 5k + 2) = 5R(k)$
- $Q_N(A_j + 5k + 3) = 5R(k+1)$
- $Q_N(A_j + 5k + 4) = 5S(k+1)$,

as required. (Of course, this assumes that $A_j$ is sufficiently large to the degree required by Proposition 7.2, but in all cases checked it appears to be sufficiently large.)

$C_j = 3$**:** The first unknown term here is $Q_N(A_j+1)$. We compute the next 4 terms, obtaining the values in the theorem statement. We observe that the sequence strongly dies because $Q_N(A_j+4) = 0$.

$C_j = 4$**:** The first unknown term here is $Q_N(A_j+4)$. We compute the next 11 terms, obtaining the values in the theorem statement. We observe that the sequence strongly dies because $Q_N(A_j+14) = 0$.

□

See Figure 9.1 for a plot of the first 30000 terms of $Q_{42}$. For $N = 42$, $j = 3$ and $C_3 = 2$, so, after the initial condition, there is the zone before it weakly dies, followed by a (very short) quasilinear piece, followed by two (successively longer) quasilinear pieces, followed by the eventual $R, S, T$-like behavior. Both axes have logarithmic scales, as otherwise the third quasilinear piece would dominate the plot. (Remember, the $A_i$'s grow very rapidly.)

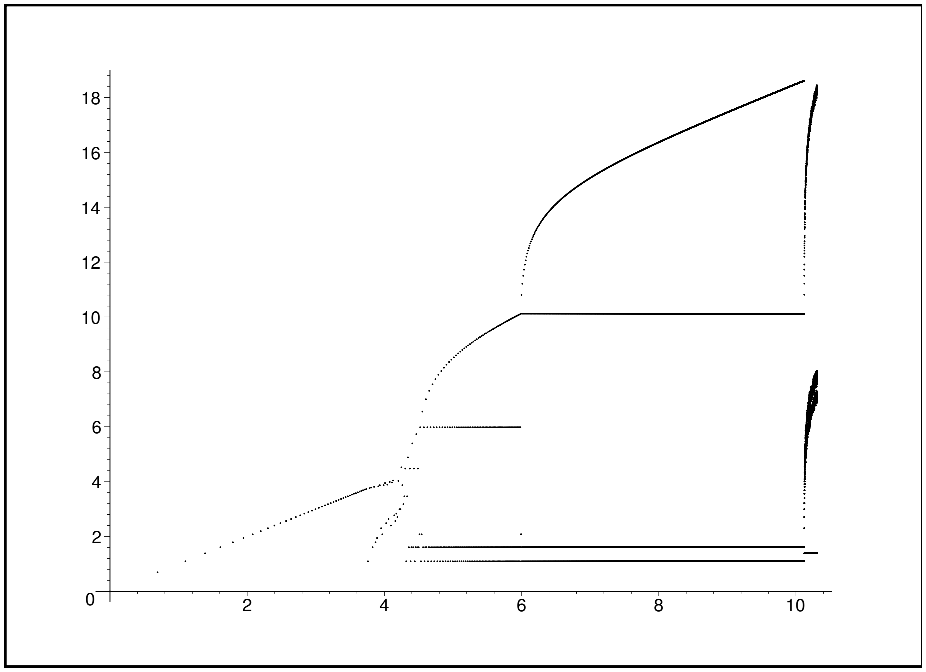


Figure 9.1: The first 30000 terms of $Q_{42}$ (A274055, both axes log scale)

Theorem 9.3 is, in a sense, rather mysterious. It completely characterizes the behavior of $Q_N$ (as long as $N$ is sufficiently large and as long as conjectures about the $R$, $S$, and $T$ sequences hold), but the characterization of which $N$ result in which behavior is somewhat convoluted. Every $N$ with $C_j < \infty$ (which is every known value of $N$) is associated to a pair $(j, C_j) \in \mathbb{Z}_{>0} \times \{0, 2, 3, 4\}$. We denote these values by $j(N)$ and $C(N)$ respectively. We also use notation like $A_i(N)$, $B_i(N)$, and $C_i(N)$ to denote $A_i$, $B_i$, and $C_i$ values for $N$. Our first observation is the following:

**Proposition 9.5.** *Let $N$ be a positive integer, and let $j = j(N)$. For all $1 \leq i \leq j$, $A_i(N + 5^j) \equiv A_i(N) \left(\operatorname{mod} 5^{j-i+1}\right)$.*

*Proof.* The proof is by induction on $i$. If $i = 1$, then $A_1(N) = 2N+4$ and $A_1(N+5^j) = 2\left(N+5^j\right)+4 = 2N+4+2\cdot 5^j$. Then, $A_1(N+5^j) - A_1(N) = 2 \cdot 5^j$, which is divisible by $5^j = 5^{j-1+1}$, as required. If $i = 2$, then

$$A_2(N) = \frac{2}{5}N^2 - 7N - \frac{78}{5}$$

and

$$A_2(N+5^j) = \frac{2}{5}N^2 - 7N - \frac{78}{5} + 2 \cdot 5^{2j-1} - 7 \cdot 5^j + 4 \cdot 5^{j-1}.$$

The difference is divisible by $5^{j-1}$, as required.

Now, suppose $i \geq 3$ and suppose that Proposition 9.5 holds for all smaller $i$ values. Recall that

$$A_i = A_{i-1}\left(\frac{A_{i-1} - A_{i-2} + 2}{5}\right) + B_{i-1}.$$

Since $i \geq 3$, $B_{i-1} = A_{i-1} - A_{i-2}$, so we can eliminate $B_{i-1}$ and write

$$A_i = A_{i-1}\left(\frac{A_{i-1} - A_{i-2} + 7}{5}\right) - A_{i-2}.$$

By induction, $A_{i-1}(N + 5^j) = A_{i-1}(N) + \alpha \cdot 5^{j-i+2}$ for some integer $\alpha$. Similarly, $A_{i-2}(N + 5^j) = A_{i-2}(N) + \beta \cdot 5^{j-i+3}$ for some integer $\beta$.

We now evaluate

$$\begin{aligned}
A_i(N+5^j) - A_i(N) = {} & A_{i-1}(N+5^j)\left(\frac{A_{i-1}(N+5^j) - A_{i-2}(N+5^j) + 7}{5}\right) \\
& - A_{i-2}(N+5^j) - A_{i-1}(N)\left(\frac{A_{i-1}(N) - A_{i-2}(N) + 7}{5}\right) \\
& - A_{i-2}(N) \\
= {} & \big(A_{i-1}(N) + \alpha \cdot 5^{j-i+2}\big) \\
& \cdot \left(\frac{\big(A_{i-1}(N) + \alpha \cdot 5^{j-i+2}\big) - \big(A_{i-2}(N) + \beta \cdot 5^{j-i+3}\big) + 7}{5}\right) \\
& - \big(A_{i-2}(N) + \beta \cdot 5^{j-i+2}\big) \\
& - A_{i-1}(N)\left(\frac{A_{i-1}(N) - A_{i-2}(N) + 7}{5}\right) - A_{i-2}(N).
\end{aligned}$$

Evaluating this expression with Maple yields

$$\begin{aligned}
A_i(N+5^j) - A_i(N) = 5^{j-i+1}\big(& 2\alpha A_{i-1}(N) - 5\beta A_{i-1}(N) - \alpha A_{i-2}(N) \\
& + \alpha^2 \cdot 5^{j-i+2} - \alpha \cdot 5^{j-i+3} + 7\alpha - 25\beta\big),
\end{aligned}$$

which is divisible by $5^{j-i+1}$, as required. $\square$

Of course, Proposition 9.5 immediately generalizes to replacing $5^j$ with any integer multiple of $5^j$. We have the following corollary to Proposition 9.5 (which also generalizes in this way):

**Corollary 9.6.** *For all $N$, and for all $1 \leq i \leq j(N)$, $C_i\big(N + 5^{j(N)}\big) = C_i(N)$. In particular, $j\big(N + 5^{j(N)}\big) = j(N)$.*

*Proof.* Let $j = j(N)$. Let $1 \le i \le j$. By Proposition 9.5,

$$A_i(N + 5^j) \equiv A_i(N) \left(\mathrm{mod}\, 5^{j-i+1}\right).$$

Since $C_i$ is a function solely of $A_i$ mod 5 and of $i$, we have $C_i(N+5^j) = C_i(N)$. Since $i$ was arbitrary, we have $C_i(N+5^j) = C_i(N)$ for every such $i$, as required. Also, $C_i(N) = 1$ if $i < j$ (by the definition of $j$). So, by the definition of $j$, we have $j\left(N + 5^j\right) = j(N)$, as required. □

Corollary 9.6 tells us that, to determine the behavior of $Q_N$, we should first look at $N$ mod 5. If $C_1(N) = 1$, then we need to look at $N$ mod 25. If $C_2(N) = 1$, then we need to look at $N$ mod 125, etc. This observation suggests that treating $N$ as a 5-adic integer may be helpful. In fact, a subset of the strings $\{0, 1, 2, 3, 4\}^*$ form a tree as follows:

- The root is the empty string, and it has the five length-1 strings as children.
- For a string $w$, interpret it as a base 5 integer $N_w$. Let $C = C_{|w|}(N_w)$ (where $|w|$ denotes the length of $w$). If $C = 1$, then $w$ has children $\{xw : x \in \{0, 1, 2, 3, 4\}\}$; otherwise $w$ is a leaf of type $C$.

To determine the behavior of $Q_N$, read the base-5 digits of $N$ from right to left and traverse the tree. When a leaf is reached, stop, and the leaf's type will determine the behavior. The tree has a level structure; level $i$ consists of the strings of length $i$ that appear in the tree.

So, the key to understanding $Q_N$ is to understand this tree (Figure 9.2). If $i \in \{1, 2, 3, 4, 5\}$, there is one leaf of each type of level $i$, as well as one internal node. (The internal nodes correspond to 2 mod 5 (2), 17 mod 25 (32), 117 mod 125 (432), 492 mod 625 (3432), and 1742 mod 3125 (23432).) It would be convenient if this structure continued, but, alas, it does not. There is no $N$ such that $j(N) = 6$; whenever $C_5(N) = 1$, $C_6(N) = 1$ also. Thereafter, it appears that all of the children of a given node are of the same type (type 0, type 2, type 3, type 4, or internal). Going deeper, there are also no $N$ with $j(N) = 11$, and there are no $N$ with $j(N) = 16$. The obvious conjecture is that there is no $N$ with $j(N) \equiv 1 \,(\mathrm{mod}\, 5)$ (except, of course, for $j = 1$).

But, the node types and locations in this tree, despite being somehow "balanced" between the five types, appear somewhat random. The main structural observation is that there appears to be a fivefold explosion of internal nodes at levels congruent to 1 mod 5, and then a twofold multiplication at the next level. Unfortunately, this exponential growth makes computing the full tree difficult beyond level 16.

Figure 9.2 depicts levels 0 through 11 of the tree. The nodes are color coded. Type 0 nodes are red, type 2 are yellow, type 3 are green, type 4 are blue, and internal (type 1) nodes are black. Figure 9.3 depicts levels 9 through 14 starting from the internal node 313223432 (the higher of the two internal nodes on level 9 with internal children in Figure 9.2). Figure 9.3 also depicts levels 13 through 17 starting from the internal node 2310313223432. (In the left part of Figure 9.3, first find the lower group of internal nodes on level 13 with internal children. Node 2310313223432 is the middle node of this group.)

We do not know if this tree is infinite. But, it has finite branching (each node has zero or five children), so, if it is infinite, König's Tree Lemma [26] implies the existence of an infinite branch. This infinite branch corresponds to a 5-adic integer $N_0$ for which $Q_N$ consists entirely of period-5 pieces, and, hence, persists forever without entering the realm of the $R$, $S$, and $T$ sequences. (See 9.2.2 for some finitary sequences with similar sorts of behaviors.) It is theoretically possible that such an $N_0$ could be an ordinary integer, though this seems unlikely.

### Analysis of Sporadic $N$ Values

Theorem 9.3 characterizes the behavior of $Q_N$ for all

$$N \notin \{n : 2 \leq n \leq 34\} \cup \{n : 1 < n < 118 \text{ and } n \equiv 1 \pmod{5}\} \cup \{57, 67, 82, 107, 117\}.$$

These 55 values can be studied individually by generating the sequences and observing the terms. A file detailing all of these findings can be found at `http://github.com/nhf216/thesis/Hof1thruN.txt`. What follows is a summary of the findings. If $N \leq 27$, $Q_N$ appears to behave chaotically and persist for a long time (at least 10 million terms), unless $N \in \{19, 23, 26\}$, in which case $Q_N$ strongly dies after not too long. Thereafter,

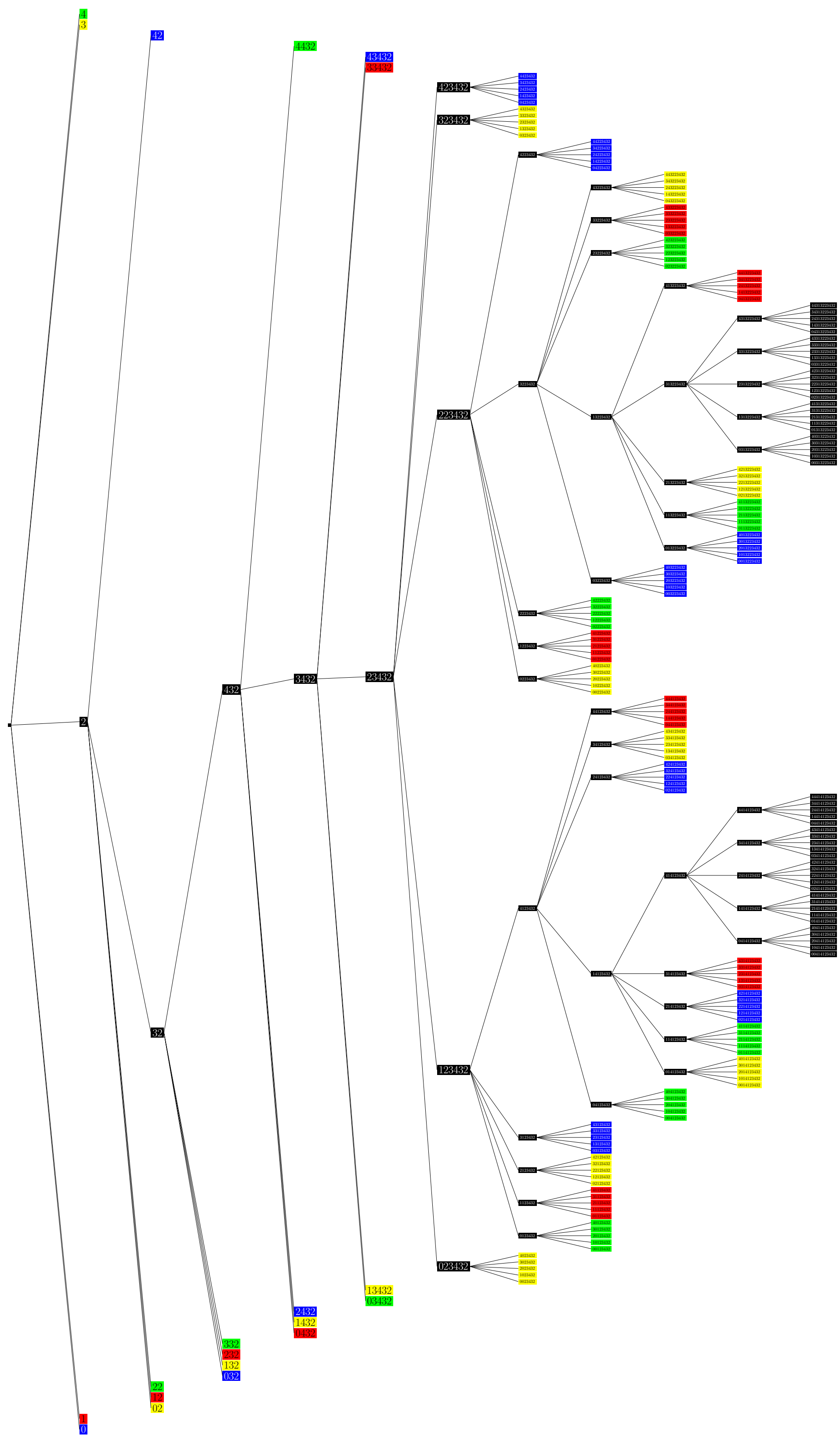

Figure 9.2: Levels 0 through 11 of the tree of behaviors

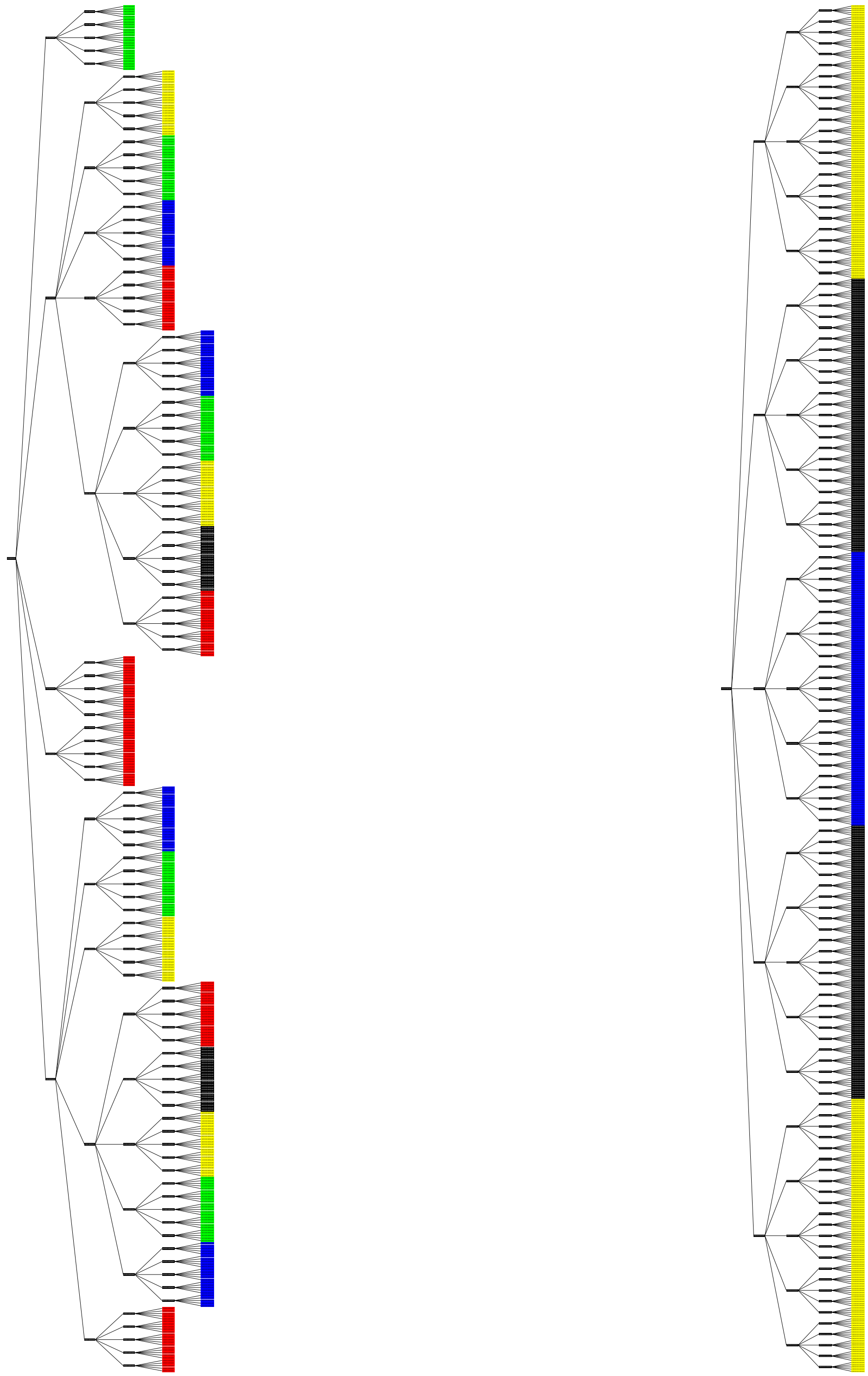

Figure 9.3: Levels 9 through 14 of the tree of behaviors, starting from 313223432 (left), and levels 13 through 17 of the tree of behaviors, starting from 2310313223432 (right)

the remaining $QN$ die, except for $N \in \{33, 36, 67, 71\}$. These all eventually behave like the case $C_j = 2$ in Theorem 9.3 and persist forever if the $R$, $S$, and $T$ sequences live forever.

Of the $N$ values exceeding 27, all but $N = 67$ and $N = 117$ can be computed explicitly until the $(R, S, T)$-pattern is entered or the sequence dies. For $N = 117$, $j = 4$, $C_4 = 0$, and Theorem 9.3 describes the first 3346939303913 terms. After this, the terms in Appendix D through $Q_{117}(A_4+156)$ are all still valid. Then, $Q_{117}(A_4+157) = 36$ instead of 151. But, the following two terms are still "large", so the sequence ends up dying after 3346939304071 terms [31, A283883]. For, $N = 67$, $j = 3$ and $C_3 = 0$. Theorem 9.3 characterizes the terms through $Q_{67}(309260)$. These terms, and some subsequent ones, can be generated easily enough. Starting with $Q_{67}(309403)$, the following pattern emerges:

- $Q_{67}(5k) = 19047817435$
- $Q_{67}(5k+1) = 3$
- $Q_{67}(5k+2) = 5$
- $Q_{67}(5k+3) = 19047817435k - 1178640853737358$
- $Q_{67}(5k+4) = 5$

This pattern breaks when indices exceed 19047817435, though it will last until then. This is too many terms to compute. But, $19047817435 \equiv 0 \pmod 5$. We then compute

- $Q_{67}(19047817436) = Q_{67}(1) + Q_{67}(19047817431) = 4$
- $Q_{67}(19047817437) = Q_{67}(19047817433) + Q_{67}(2) = 725626911475164410 54.$

By Proposition 7.2, this settles into a pattern like the $C_j = 2$ case of Theorem 9.3, and this pattern persists as long as the $R$, $S$, and $T$ sequences live (which we conjecture to be forever) [31, A283882].

## 9.2 Three-term Hofstadter Recurrence

In the previous section, we consider the Hofstadter $Q$-recurrence $Q(n) = Q(n - Q(n - 1)) + Q(n - Q(n - 2))$ with initial conditions $|1, 2, 3, \ldots, N|$. In this section and the next section, we consider this same initial condition as input to obvious generalizations of the $Q$-recurrence. In the current section, we examine the three-term recurrence $B(n) = B(n - B(n - 1)) + B(n - B(n - 2)) + B(n - B(n - 3))$, which we previously encountered in Chapters 6 and 3 (and also briefly near the end of Chapter 5).

We have the following definition:

**Definition 9.7.** *Let $N \geq 3$ be a positive integer. Define the sequence $B_N(n)$ to be the sequence generated by the above $B$-recurrence above with the initial condition $\langle 1, 2, 3, \ldots, N \rangle$.*

### 9.2.1 General Structure of $B_N$

In this subsection, we discuss the general behavior of the sequences $B_N$. Theorem 9.8, classifies the behavior for all but finitely many $N$. The remaining subsections analyze some specific values from these finitely many remaining $N$.

**Theorem 9.8.** *Let $N$ be a natural number. Then, $B_N$ weakly dies after $N + 24$ terms, provided $N \geq 14$. Under the strong death convention, if $N \geq 74$, then the following period-7 pattern begins at index $N + 67$:*

- $B_N(N + 7k) = 7k + 2$
- $B_N(N + 7k + 1) = N + 7k + 2$
- $B_N(N + 7k + 2) = N + 7k + 4$
- $B_N(N + 7k + 3) = 7$
- $B_N(N + 7k + 4) = 2N + 2k + 45$
- $B_N(N + 7k + 5) = 2N + k - 7$
- $B_N(N + 7k + 6) = N - 2$

*This pattern lasts through index* $2N+\nu$*, where* $\nu = \max(-2, -3 + ((2-N) \bmod 7))$*. After this,*

- *If* $N \equiv 0 \pmod 7$ *and* $N \geq 196$*, then* $B_N$ *strongly dies after* $2N+27$ *terms.*
- *If* $N \equiv 1 \pmod 7$ *and* $N \geq 2087$*, then* $B_N$ *strongly dies after* $2N+254$ *terms.*
- *If* $N \equiv 2 \pmod 7$ *and* $N \geq 3201$*, then* $B_N$ *strongly dies after* $2N+524$ *terms.*
- *If* $N \equiv 3 \pmod 7$ *and* $N \geq 4315$*, then* $B_N$ *strongly dies after* $2N+560$ *terms.*
- *If* $N \equiv 4 \pmod 7$ *and* $N \geq 200$*, then* $B_N$ *strongly dies after* $2N+20$ *terms.*
- *If* $N \equiv 5 \pmod 7$ *and* $N \geq 32478$*, then* $B_N$ *strongly dies after* $2N+4547$ *terms.*
- *If* $N \equiv 6 \pmod 7$ *and* $N \geq 118$*, then* $B_N$ *strongly dies after* $2N+9$ *terms.*

*Proof.* The proof of Theorem 9.8 is similar in style to the proof of Theorem 9.3, but much longer (due to the large constants appearing in the theorem) and somewhat more tedious (due to the need to prove a period-7 pattern instead of a period 5 pattern and the need to check thousands of assumptions on $N$ to see which ones are dominant). For this reason, we omit the full proof of Theorem 9.8 and instead only summarize it here. All of the terms of $B_N$ for general $N$ can be found in the file `http://github.com/nhf216/thesis/trihofform.txt`. These were generated by the procedure `ProveLongTermEventualSolution` in `http://github.com/nhf216/thesis/nonstdhof.txt`. There are seven items in this file, one for each congruence class for $N$ mod 7.

The first thing to do is to generate generic terms of $B_N$, starting from $B_N(N+1)$. Doing so, we observe that $B_N(N+24) = 2N+11$. Generating this sequence to this point requires that $N$ be at least 9. If $N \geq 14$, then $2N+11 \geq N+25$, so, in that case, the sequence weakly dies after $N+24$ terms, as required.

Continuing to generate terms using the strong death convention, we notice that the proposed period-7 pattern develops beginning in index $N+67$, assuming $N \geq 74$. Proving that this pattern persists for awhile follows from a straightforward but tedious inductive argument akin to the one used in Lemma 9.4. We then need to determine

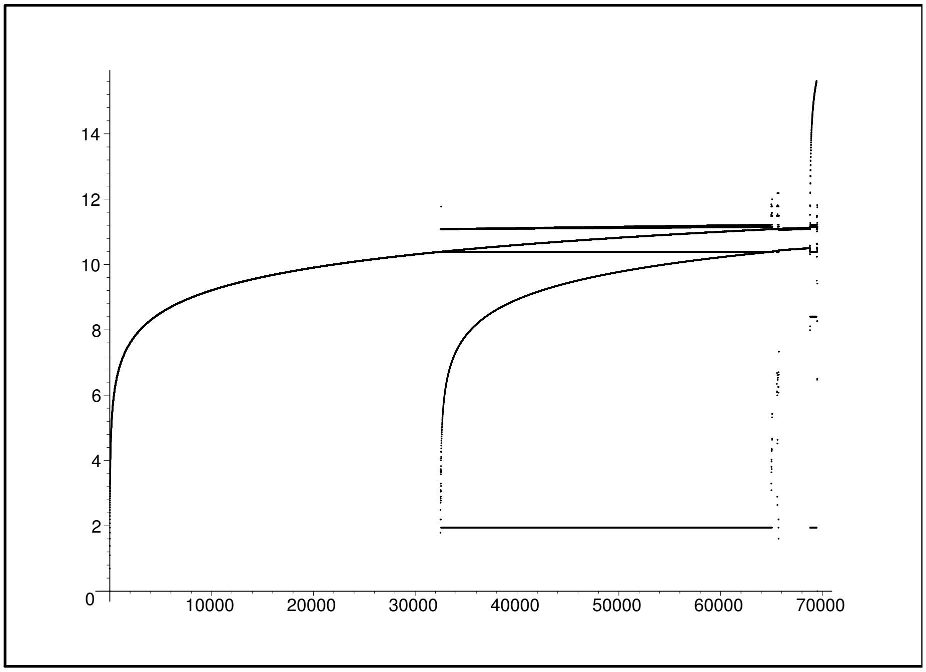


Figure 9.4: All 69503 terms of $B_{32478}$ (A274058, log plot)

how long the pattern persists. The only cases of the inductive argument that make any assumptions that eventually are violated are the ones that refer to $B_N(N+7k+6)$ case. We assume inductively in these three cases (the $N+7k$, $N+7k+1$, and $N+7k+2$ cases) that subtracting $N-2$ puts us into the initial condition. So, if $(2N-1 \bmod 7) \in \{0,1,2\}$, then the pattern ceases after index $2N-2$. Otherwise, the pattern continues until we reach an index that is divisible by 7, at which point it ceases. This behavior is summarized by the constant $\nu$ defined in the theorem statement.

The final step is to generate more terms after the pattern ends. This, of course, depends on $N \bmod 7$, so there are seven cases to consider. In each case, the sequence strongly dies after a constant number of additional terms. When computing these terms, we keep track of the assumptions we make on the size of $N$. Doing both of these things results in the totality of the final part of the theorem statement. $\square$

Figure 9.4 shows the full sequence $B_{32478}$. The $y$-axis is logarithmic, as otherwise only the giant terms near the end would be visible.

Theorem 9.8 describes the behavior of $B_N$ for all but 6079 values of $N$. A file containing information about these unclassified values can be found at `http://github.com/nhf216/thesis/TriHof1thruN.txt`. What follows is a summary of the findings. We know that $B_5$ and $B_6$ live forever (do not weakly die), as these slow sequences are the subject of Chapter 6. (These are actually the same sequence.) If $N \in \{7, 8, 9\}$, $B_N$ is not known to die weakly. (These sequence each live for at least $10^8$ terms.) For all other values of $N$, $B_N$ dies weakly. For

$$N \in \{5, 6, 81, 182, 193, 429, 822, 1892, 2789, 3442, 7292, 23511, 25163\},$$

$B_N$ is known not to strongly die. For $N \in \{4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 18\}$, it is unknown whether $B_N$ strongly dies. For all other $N$, $B_N$ strongly dies, though $B_{20830}$ lives for $84975 \cdot 2^{560362} + 31$ terms.

The following proposition leads to a proof that $B_N$ does not strongly die for $N \in \{81, 182, 429, 822, 1892, 2789, 7292, 23511, 25163\}$.

**Proposition 9.9.** *Let $K$ be a nonnegative integer, and let $M \geq K + 5$. For any integers $a_1, a_2, \ldots, a_K$, the initial condition $\langle a_1, a_2, \ldots, a_K, 2, M, 2\rangle$ to the $B$-recurrence generates a sequence with pattern*

- $B_p(K + 2k) = 2^{k-1} \cdot M$
- $B_p(K + 2k + 1) = 2$

*for all indices greater than $K$.*

*Proof.* The proof is by induction on the index. The indices $K + 1$, $K + 2$, and $K + 3$ constitute the base case, and the proposition holds for these as they fall in the initial condition.

Now, suppose the index is at least $K + 4$, and suppose the proposition holds for all prior indices. There are two cases to consider.

**Index $K+2k$:** In this case, we have

$$\begin{aligned}B_p(K+2k) &= B_p(K+2k-B_p(K+2k-1)) + B_p(K+2k-B_p(K+2k-2)) \\ &\quad + B_p(K+2k-B_p(K+2k-3)) \\ &= B_p(K+2k-2) + B_p(K+2k-2^{k-2}\cdot M) + B_p(K+2k-2) \\ &= 2^{k-2}\cdot M + B_p(K+2k-2^{k-2}\cdot M) + 2^{k-2}\cdot M.\end{aligned}$$

Since $M \geq K+5$ and $k \geq 2$, we have that $B_p(K+2k-2^{k-2}\cdot M) = 0$, so $B_p(K+2k) = 2^{k-1}\cdot M$, as required.

**Index $K+2k+1$:** In this case, we have

$$\begin{aligned}B_p(K+2k+1) &= B_p(K+2k+1-B_p(K+2k)) \\ &\quad + B_p(K+2k+1-B_p(K+2k-1)) \\ &\quad + B_p(K+2k+1-B_p(K+2k-2)) \\ &= B_p(K+2k+1-2^{k-1}\cdot M) + B_p(K+2k+1-2) \\ &\quad + B_p(K+2k+1-2^{k-2}\cdot M) \\ &= B_p(K+2k+1-2^{k-1}\cdot M) + 2 + B_p(K+2k+1-2^{k-2}\cdot M).\end{aligned}$$

Since $M \geq K+5$ and $k \geq 2$, we have that $B_p(K+2k+1-2^{k-1}\cdot M) = 0$ and $B_p(K+2k+1-2^{k-2}\cdot M) = 0$, so $B_p(K+2k+1) = 2$, as required.

□

When computing terms of $B_N$ for

$$N \in \{81, 182, 429, 822, 1892, 2789, 7292, 23511, 25163\},$$

eventually a consecutive subsequence $2, M, 2$ appears for some sufficiently large $M$. (In fact, usually the occurrence of a 2 late in the sequence is enough to cause this to happen.) Then, by Proposition 9.9, the sequence persists with this alternating pattern. Philosophically, it feels somewhat like these sequences strongly die, but they eventually contain a 2 rather than a 0. For this reason, we might say that these sequences have "lobotomies," but they do not strongly die.

| N | K | M |
|---|---|---|
| 81 | 527 | 565 |
| 182 | 390 | 461 |
| 429 | 1268 | 1313 |
| 822 | 3689 | 4161 |
| 1892 | 5872 | 6103 |
| 2789 | 9234 | 10510 |
| 7292 | 22135 | 22948 |
| 23511 | 69983 | 70559 |
| 25163 | 75661 | 82457 |

Table 9.1: $B_N$: $K$ and $M$ values for special $N$ values

The $K$ and $M$ values (as in Proposition 9.9) for each of these $N$ values are summarized in Table 9.1.

Most of the remaining strongly-dying $B_N$ can be computed entirely, right up until a zero occurs. Assuming these computations handle any sequence dying before index 80 million, only the cases $N = 193$, $N = 3442$, $N = 19395$, $N = 20830$, and $N = 27298$ remain to be analyzed. The first two of these are handled in the next subsection, the last three in Subsection 9.2.3.

### 9.2.2 Analysis of $B_{193}$ and $B_{3442}$

In this subsection, we show that the sequences $B_{193}$ (A283884) and $B_{3442}$ (A283885) persist. They do not persist because of an interaction with Proposition 9.9; instead, they have a more complicated recursive structure. To easily describe what happens in these sequences, we need the following proposition, which can be thought of as a relative of Lemma 9.4 and Proposition 9.9.

**Proposition 9.10.** *Let $K \geq 3$ and $\mu \geq 1$ be integers. Then, for any integers $a_4, a_5, \ldots, a_K$, the initial condition*

$$\langle 1, 2, 3, a_4, a_5, \ldots, a_K, K + \mu, 3, K + 3, K + \mu + 1, 5\rangle$$

*to the $B$-recurrence generate the pattern*

- $B_C(K + 5k) = 5$

- $B_C(K+5k+1) = K+3k+\mu$
- $B_C(K+5k+2) = 3$
- $B_C(K+5k+3) = K+5k+3$
- $B_C(K+5k+4) = K+3k+\mu+1$

*through* $B_C\left(K+\left\lfloor\frac{5\mu-15}{2}\right\rfloor\right)$.

*Proof.* As usual, the proof is a straightforward induction argument, so we leave the details as an exercise. When checking the different cases, the assumptions that $B_C(1) = 1$, $B_C(2)$, $B_C(3) = 3$, $3k+\mu \geq 5k+4$, and $3k+\mu+1 \geq 5k+7$ are necessary. These last two are the only assumptions that are eventually violated. This happens when the slope-1 terms overtake the slope-$\frac{3}{5}$ terms, which is first when $k = \frac{\mu}{2} - 3$. Converting this condition to an index results in the final part of the proposition, regarding the end of the pattern. □

Analysis of $B_{193}$ and of $B_{3442}$ involves Proposition 9.10. After generating some terms of $B_{193}$, we observe that it settles in to a pattern described by Proposition 9.10 with parameters $K = 441$ and $\mu = 793$. This pattern then lasts until at least index 2416; in fact, it lasts through $B_{193}(2417)$, but then $B_{193}(2418) = 4$. Later on, the sequence arrives at another such pattern with $K = 2858$ and $\mu = 5627$ that persists through $B_{193}(16919)$, after which $B_{193}(16920) = 4$. More of these patterns continue, and they are summarized by the following proposition:

**Proposition 9.11.** *Let* $K_0 = 441$*, and recursively let* $K_i = 6K_{i-1} + 212$ *for* $i \geq 1$*. The sequence* $B_{193}$*, beginning at index* 442*, consists entirely of patterns described by Proposition 9.10 with parameters* $K = K_i$ *and* $\mu = 2K_i - 89$ *for each* $i$*. Pattern* $i$ *persists until index* $6K_i - 229 = K_{i+1} - 441$*, after which there are* 441 *sporadic terms that are parametrized by* $i$*.*

*Proof.* Let $B_{193,K}$ denote the sequence generated by the $B$-recurrence with initial condition

$$\langle B_{193}(1), B_{193}(2), \ldots, B_{193}(440), B_{193}(441), a_{442}, a_{443}, \ldots, a_{K-1}, a_K, \\ 3K-89, 3, K+3, 3K-88, 5\rangle\,,$$

where $a_{442}$ through $a_K$ are arbitrary integers. We claim the following about $B_{193,K}$:

1. A Proposition 9.10 pattern with $K = K$ and $\mu = 2K-89$ persists from index $K+1$ through index $6K-229$.

2. This sequence is an instance of $B_{193,6K+212}$. (By *instance*, we mean that some of the $a_j$ values have been assigned specific values.)

Item 1 follows almost immediately from Proposition 9.10. The only thing that needs to be checked is the termination point, which is easy to verify.

To prove item 2, we use the procedure `KExplore193` in `http://github.com/nhf216/thesis/ProveTriHof1thruN.txt` to generate terms and find and prove the next pattern that $B_{193,K}$ reaches. Provided $K \geq 111$, this pattern is precisely a Proposition 9.10 pattern with $K = 6K+212$ and $\mu = 12K+335$. Since the form of the general sequence $B_{193,K}$ is completely determined by the first 441 and final five terms of the initial condition, it is clear that $B_{193,K}$ is an instance of $B_{193,6K+212}$, as required.

Since $B_{193}$ itself is $B_{193,K_0}$ (and $K_0 \geq 111$), Proposition 9.11 follows. □

For a listing of the 441 sporadic terms, see `http://github.com/nhf216/thesis/TriHof193Sporadic.txt`. Interestingly, the $i^{th}$ instance of these terms contain a brief Proposition 9.10 pattern with $K = 6K_i + 54$ and $\mu = 54$. For a plot of $B_{193}$, see Figure 9.5.

The behavior of $B_{3442}$ is similar to that of $B_{193}$, and it is governed by a similar rule to Proposition 9.11.

**Proposition 9.12.** *Let* $K_0 = 95123$*, and recursively thereafter let* $K_i = 6K_{i-1} + 11714$ *for* $i \geq 1$*. The sequence* $B_{3442}$*, beginning at index* 95124*, consists entirely of patterns described by Proposition 9.10 with parameters* $K = K_i$ *and* $\mu = 2K_i - 89$ *for each* $i \geq 1$*.*

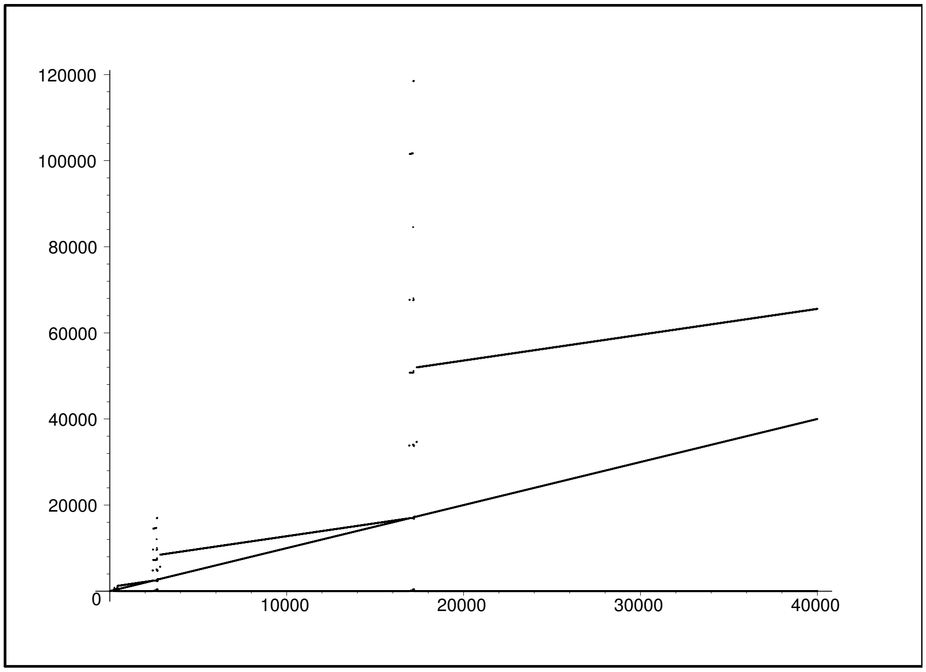


Figure 9.5: First 40000 terms of $B_{193}$ (A283884)

*Pattern $i$ persists until index* $6K_i - 229 = K_{i+1} - 11943$*, after which there are* 11943 *sporadic terms that are parametrized by $i$.*

The proof of Proposition 9.12 is nearly identical to that of Proposition 9.11. Instead of $K_0 \geq 111$, we need $K_0 \geq 1457$, but we do have that. The proof is carried out by the procedure `KExplore193` in `http://github.com/nhf216/thesis/ProveTriHof1thruN.txt`. For a listing of the 11943 sporadic terms, see `http://github.com/nhf216/thesis/TriHof3442Sporadic.txt`. Much like $B_{193}$, the $i^{th}$ instance of these terms in $B_{3442}$ contains a brief Proposition 9.10 pattern with $K = 6K_i + 3286$ and $\mu = 3362$. Also, $B_{3442}$ has an earlier, long Proposition 9.10 pattern with $K = 13889$ and $\mu = 27719$ that lasts through index 83180.

Other sequences besides $B_{193}$ and $B_{3442}$ contain Proposition 9.10 patterns but die afterwards. Examples include $B_{20592}$, $B_{23378}$, and $B_{32471}$.

### 9.2.3 Analysis of $B_{19395}$, $B_{20830}$, and $B_{27298}$

In this subsection, we study the sequences $B_{19395}$, $B_{20830}$, and $B_{27298}$ (A283886, A283887, and A283888 respectively). These three sequences all strongly die, but they all last too long to reasonably compute all of the terms, and the terms become large very rapidly. ($B_{19395}$ dies after 80444792 terms, $B_{20830}$ dies after $84975 \cdot 2^{560362} + 31$ terms, and $B_{27298}$ dies after 141895479 terms.) Even though they all strongly die, the analysis of these sequences will resemble the analysis of $B_{193}$ and $B_{3442}$ from Subsection 9.2.2. Much like before, we have a unifying pattern that will assist us, though this time the pattern is somewhat more complicated.

**Proposition 9.13.** *Let $K$ be a nonnegative integer. Let $\lambda$, $\mu_1$, $\mu_2$, and $\nu$ be any positive integers with $\lambda \geq 31 + K$ and the others exceeding $\lambda$. Then, for any integers $a_1, a_2, \ldots, a_K$ the initial condition*

$$\langle a_1, a_2, \ldots, a_K, \lambda, 7, \mu_2, 16, \mu_2, 16, \mu_1, \lambda, 7, \mu_2, 16, 2\mu_2, 16, \mu_2, 25, \nu, \lambda, 7 \rangle$$

*to the $B$-recurrence generates the pattern*

- $B_U(K + 16k) = \mu_1 \cdot 2^{k-1} + \nu - \mu_1$
- $B_U(K + 16k + 1) = \lambda$
- $B_U(K + 16k + 2) = 7$
- $B_U(K + 16k + 3) = \mu_2 \cdot 2^k$
- $B_U(K + 16k + 4) = 16$
- $B_U(K + 16k + 5) = \mu_2 \cdot 2^k$
- $B_U(K + 16k + 6) = 16$
- $B_U(K + 16k + 7) = \mu_1 \cdot 2^k$
- $B_U(K + 16k + 8) = \lambda$
- $B_U(K + 16k + 9) = 7$
- $B_U(K + 16k + 10) = \mu_2 \cdot 2^k$
- $B_U(K + 16k + 11) = 16$
- $B_U(K + 16k + 12) = \mu_2 \cdot 2^{k+1}$
- $B_U(K + 16k + 13) = 16$
- $B_U(K + 16k + 14) = \mu_2 \cdot 2^k$
- $B_U(K + 16k + 15) = 25$

*through term $B_U(\lambda)$. (The pattern begins in index $K + 1$.)*

*Proof.* As usual, the proof is a fairly straightforward but tedious induction argument, so we leave the details as an exercise. The only somewhat interesting case is $K + 16k$, so we do that one here explicitly.

$$\begin{aligned}
B_U(K+16k) &= B_U(K+16k-B_U(K+16k-1)) \\
&\quad + B_U(K+16k-B_U(K+16k-2)) \\
&\quad + B_U(K+16k-B_U(K+16k-3)) \\
&= B_U(K+16k-25) + B_U(K+16k-\mu_2\cdot 2^{k-1}) \\
&\quad + B_U(K+16k-16) \\
&= \mu_1\cdot 2^{k-2} + 0 + \mu_1\cdot 2^{k-2} + \nu - \mu_1 \\
&= \mu_1\cdot 2^{k-1} + \nu - \mu_1,
\end{aligned}$$

as required. (Automatically discovering the formula for this case requires solving a nonhomogenous linear recurrence, but checking the formula is quite mechanical.)

When checking the different cases, the assumptions that $\lambda$ and each of the non-constant terms are always at least the index are necessary. Only the first of these assumptions is eventually violated, which happens after index $\lambda$, as required. □

(The parameter bounds in this proposition are not tight, but they suffice for our purposes.)

We begin by analyzing $B_{19395}$. We claim that this sequence strongly dies after 80444792 terms. To obtain this result, we make use of the following lemma, which is related to the proofs of Propositions 9.11 and 9.12.

**Lemma 9.14.** *Let $\lambda \geq 78116$ be an integer congruent to 4 mod 16. Let $B_{19395,\lambda}$ denote the sequence generated by the $B$-recurrence with initial condition*

$$\langle B_{19395}(1), B_{19395}(2), \ldots, B_{19395}(77733), B_{19395}(77734), \lambda, 7, 310800, 16, 310800,$$
$$16, 321833900, \lambda, 7, 310800, 16, 621600, 16, 310800, 25, 402278561, \lambda, 7\rangle\,.$$

*We claim the following about $B_{19395,\lambda}$:*

1. *A Proposition 9.13 pattern with $K = 77734$, $\lambda = \lambda$, $\mu_1 = 321833900$, $\mu_2 = 310800$, and $\nu = 402278561$ persists from index 77735 through index $\lambda + 1$.*

2. *This sequence strongly dies after $\lambda + 132$ terms, because $B_{19395,\lambda}(\lambda + 132) = 0$.*

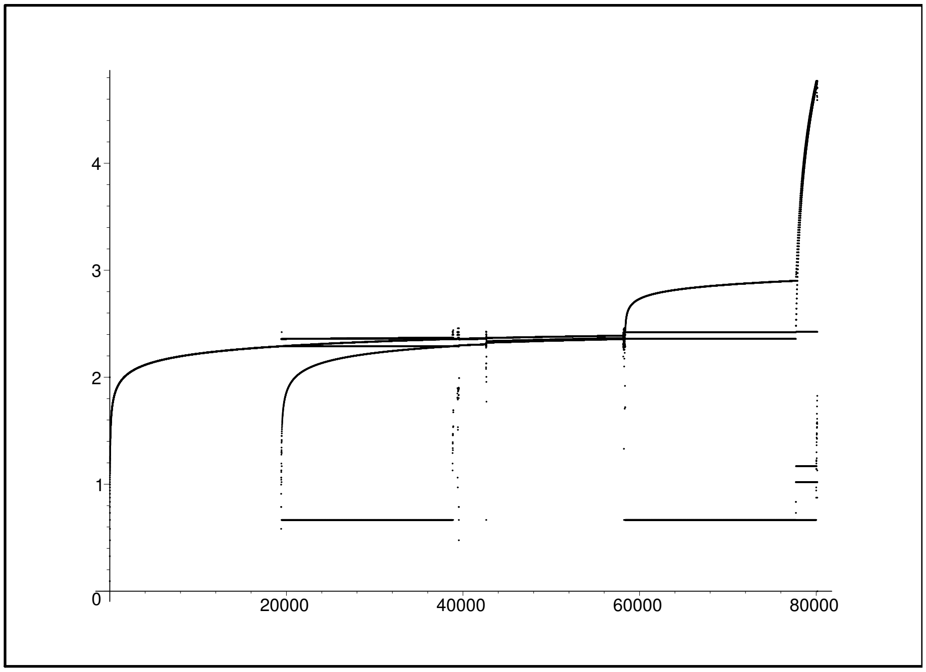


Figure 9.6: All 80136 terms of $B_{19395,80004}$ (A283886, $y$-axis doubly logarithmic)

*Proof.* Item 1 follows almost immediately from Proposition 9.13 and the definition of $B_{19395,\lambda}$. The only extra thing to check is that $B_{19395,\lambda}(\lambda+1)$ is what it is supposed to be, which can easily be verified. Item 2 can also be easily checked by a computer. The computation is only valid if $\lambda \geq 78110$, and 78116 is the next number congruent to 4 mod 16 exceeding 78110. This proof is carried out by the procedure `Explore19395` in `http://github.com/nhf216/thesis/ProveTriHof1thruN.txt`. For a listing of the 131 final terms, see `http://github.com/nhf216/thesis/TriHof19395Lfinal.txt` or `http://github.com/nhf216/thesis/TriHof19395final.txt`. □

The fact that $B_{19395}$ strongly dies after 80444792 terms follows from Lemma 9.14 and from the facts that $B_{19395} = B_{19395,80444660}$, $80444660 \equiv 4 \pmod{16}$, and $80444660 \geq 78116$.

We cannot give a plot of $B_{19395}$, but we can give a plot of $B_{19395,80004}$. Such a plot is given in Figure 9.6. The $y$-axis is doubly logarithmic.

We now analyze $B_{27298}$. We claim that this sequence strongly dies after 141895479 terms. Here, we make use of the following lemma, which is quite similar to Lemma 9.14.

**Lemma 9.15.** *Let $\lambda \geq 113441$ be an integer congruent to $1 \bmod 16$. Let $B_{27298,\lambda}$ denote the sequence generated by the $B$-recurrence with initial condition*

$$\langle B_{27298}(1), B_{27298}(2), \ldots, B_{27298}(112948), B_{27298}(112949), \lambda, 7, 903192, 16, 903192, 16, 1135082696, \lambda, 7, 903192, 16, 1806384, 16, 903192, 25, 1276977978, \lambda, 7\rangle .$$

*We claim the following about $B_{27298,\lambda}$:*

1. *A Proposition 9.13 pattern with $K = 112949$, $\lambda = \lambda$, $\mu_1 = 1135082696$, $\mu_2 = 903192$, and $\nu = 1276977978$ persists from index $112950$ through index $\lambda + 1$.*

2. *This sequence strongly dies after $\lambda + 198$ terms, because $B_{27298,\lambda}(\lambda + 198) = 0$.*

*Proof.* Item 1 follows almost immediately from Proposition 9.13 and the definition of $B_{27298,\lambda}$. The only extra thing to check is that $B_{27298,\lambda}(\lambda+1)$ is what it is supposed to be, which can easily be verified. Item 2 can also be easily checked by a computer. The computation is only valid if $\lambda \geq 113438$, and 113441 is the next number congruent to 1 mod 16 exceeding 113438. This proof is carried out by the procedure `Explore27298` in `http://github.com/nhf216/thesis/ProveTriHof1thruN.txt`. For a listing of the 197 final terms, see `http://github.com/nhf216/thesis/TriHof27298Lfinal.txt` or `http://github.com/nhf216/thesis/TriHof27298final.txt`. □

The fact that $B_{27298}$ strongly dies after 141895479 terms follows from Lemma 9.15 and from the facts that $B_{27298} = B_{27298,141895281}$, $141895281 \equiv 1 \pmod{16}$, and $141895281 \geq 113441$.

We conclude with a study of $B_{20830}$, the most complicated of the three sequences in this subsection. We claim that this sequence strongly dies after $84975 \cdot 2^{560362} + 31$ terms. Whereas the preceding two sequences could conceivably be computed in their entireties (though the terms of those sequences become quite large), $B_{20830}$ does not stand a chance to be computed fully. There are two intermediate claims on our way to describing this sequence. First, we need to describe another pattern.

**Proposition 9.16.** *Let $K$ be a nonnegative integer. Let $\lambda$, $\mu_1$, $\mu_2$, $\nu_1$, $\nu_2$, and $\nu_3$ be any positive integers with $\lambda \geq 31+K$ and the others exceeding $\lambda$. Then, for any integers*

*$a_1, a_2, \ldots, a_K$ the initial condition*

$$\langle a_1, a_2, \ldots, a_K, 16, \mu_2, 7, \nu_2, \lambda, 16, \lambda, 16, \mu_1, 10, \nu_3, \mu_2, 7, \lambda, 16, \nu_1 \rangle$$

*to the $B$-recurrence generates the pattern*

- $B_{U'}(K+16k) = \nu_1 + \left(2^k - 2\right)\mu_1$
- $B_{U'}(K+16k+1) = 16$
- $B_{U'}(K+16k+2) = \mu_2 \cdot 2^k$
- $B_{U'}(K+16k+3) = 7$
- $B_{U'}(K+16k+4) = 7k + \nu_2$
- $B_{U'}(K+16k+5) = \lambda$
- $B_{U'}(K+16k+6) = 16$
- $B_{U'}(K+16k+7) = \lambda$
- $B_{U'}(K+16k+8) = 16$
- $B_{U'}(K+16k+9) = \mu_1 \cdot 2^k$
- $B_{U'}(K+16k+10) = 10$
- $B_{U'}(K+16k+11) = 16k + \nu_3$
- $B_{U'}(K+16k+12) = \mu_2 \cdot 2^k$
- $B_{U'}(K+16k+13) = 7$
- $B_{U'}(K+16k+14) = \lambda$
- $B_{U'}(K+16k+15) = 16$

*through term $B_{U'}(\lambda)$. (The pattern begins in index $K+1$.)*

*Proof.* As usual, the proof is a fairly straightforward but tedious induction argument, so we leave the details as an exercise. The only somewhat interesting case is $K+16k$, so we do that one here explicitly.

$$\begin{aligned}
B_{U'}(K+16k) &= B_{U'}(K+16k - B_{U'}(K+16k-1)) \\
&\quad + B_{U'}(K+16k - B_{U'}(K+16k-2)) \\
&\quad + B_{U'}(K+16k - B_{U'}(K+16k-3)) \\
&= B_{U'}(K+16k-16) + B_{U'}(K+16k-\lambda) \\
&\quad + B_{U'}(K+16k-7) \\
&= \nu_1 + \left(2^{k-1} - 2\right)\mu_1 + 0 + \mu_1 \cdot 2^{k-1} \\
&= \nu_1 + \left(2^k - 2\right)\mu_1,
\end{aligned}$$

as required.

When checking the different cases, the assumptions that $\lambda$ and each of the non-constant terms are always at least the index are necessary. Only the first of these assumptions is eventually violated, which happens after index $\lambda$, as required. $\square$

(The parameter bounds in this proposition are not tight, but they suffice for our purposes.)

We now introduce a fairly complicated lemma:

**Lemma 9.17.** *Let $\lambda \geq 85031$ be an integer congruent to $7 \bmod 16$. Let $B_{20830,\lambda}$ denote the sequence generated by the $B$-recurrence with initial condition*

$$\langle B_{20830}(1), B_{20830}(2), \ldots, B_{20830}(85008), B_{20830}(85009), \lambda, 7, 339900, 16, 339900, 16, 36128364, \lambda, 7, 339900, 16, 679800, 16, 339900, 25, 45179140, \lambda, 7\rangle .$$

*Next, let $\lambda' \geq \lambda + 125$ be an integer congruent to $0 \bmod 16$. (The constant $125$ is not tight.) Select integers $\nu'_2 < \mu'_2 < \mu'_1 < \nu'_1 < \nu'_3$ with $\nu'_2 > \lambda'$. Let $B_{20830,\lambda',\mu'_1,\mu'_2,\nu'_1,\nu'_2,\nu'_3}$ denote the sequence generated by the $B$-recurrence with initial condition*

$$\Big\langle B_{20830,\lambda}(1), B_{20830,\lambda}(2), \ldots, B_{20830,\lambda}(\lambda+24), B_{20830,\lambda}(\lambda+25), 16, \mu'_2, 7, \nu'_2, \lambda', 16, \lambda', 16, \mu'_1, 10, \nu'_3, \mu'_2, 7, \lambda', 16, \nu'_1\Big\rangle .$$

*We claim the following about $B_{20830,\lambda}$ and about $B_{20830,\lambda',\mu'_1,\mu'_2,\nu'_1,\nu'_2,\nu'_3}$:*

1. *A Proposition 9.13 pattern with $K = 85009$, $\lambda = \lambda$, $\mu_1 = 36128364$, $\mu_2 = 339900$, and $\nu = 45179140$ persists from index $85010$ through index $\lambda$.*

2. *The sequence $B_{20830,\lambda}$ is the same as the sequence $B_{20830,\lambda',\mu'_1,\mu'_2,\nu'_1,\nu'_2,\nu'_3}$ with*

   - $\lambda' = \mu_2 \cdot 2^{560360}$
   - $\mu'_1 = 2\lambda + \mu_2 \cdot 2^{560361}$
   - $\mu'_2 = 2\lambda + \mu_2 \cdot 2^{560360}$
   - $\nu'_1 = 2\mu_1 + \mu_2 \cdot 2^{560362}$
   - $\nu'_2 = 15 + \mu_2 \cdot 2^{560360}$
   - $\nu'_3 = \lambda + 35 + \mu_1 \cdot 2^{560359}$.

*(In particular,* $\lambda' \equiv 0 \pmod{16}$*, and we have* $\lambda' < \nu_2' < \mu_2' < \mu_1' < \nu_1' < \nu_3'$*.)*

3. *A Proposition 9.16 pattern with* $K = \lambda + 25$*,* $\lambda = \lambda'$*,* $\mu_1 = \mu_1'$*,* $\mu_2 = \mu_2'$*,* $\nu_1 = \nu_1'$*,* $\nu_2 = \nu_2'$*, and* $\nu_3 = \nu_3'$ *persists from index* $\lambda + 26$ *through index* $\lambda'$*.*

4. *The sequence* $B_{20830,\lambda',\mu_1',\mu_2',\nu_1',\nu_2',\nu_3'}$ *strongly dies after* $\lambda' + 31$ *terms, because* $B_{20830,\lambda',\mu_1',\mu_2',\nu_1',\nu_2',\nu_3'}(\lambda' + 31) = 0$.

*Proof.* Item 1 follows immediately from Proposition 9.13 and the definition of $B_{20380,\lambda}$. Item 2 can be easily checked by a computer. Generating 41 terms beginning with $B_{20380,\lambda}(\lambda + 1)$ yields terms that form an initial condition in line with the definition of $B_{20830,\lambda',\mu_1',\mu_2',\nu_1',\nu_2',\nu_3'}$. The computations are only valid if $\lambda \geq 85025$, and 85031 is the next number congruent to 7 mod 16 exceeding 85025.

Item 3 follows immediately from Proposition 9.16 and the definition of $B_{20830,\lambda',\mu_1',\mu_2',\nu_1',\nu_2',\nu_3'}$. Item 4 can be easily checked by a computer. Generating 31 terms beginning with $B_{20830,\lambda',\mu_1',\mu_2',\nu_1',\nu_2',\nu_3'}(\lambda' + 1)$ yields 31 terms, where the last one is 0.

This proof is carried out by the procedures `Explore20380a` and `Explore20380c` in `http://github.com/nhf216/thesis/ProveTriHof1thruN.txt`. For a listing of the intermediate 25 terms between the two patterns, see `http://github.com/nhf216/thesis/TriHof20830Lmid.txt` or `http://github.com/nhf216/thesis/TriHof20830mid.txt`. For a listing of the final 31 terms, see `http://github.com/nhf216/thesis/TriHof20830Lfinal.txt` or `http://github.com/nhf216/thesis/TriHof20830final.txt`. ☐

The fact that $B_{20830}$ strongly dies after $84975 \cdot 2^{560362} + 31$ terms follows from Lemma 9.17 and from the following facts:

- $B_{20830} = B_{20830,9050775}$
- $9050775 \equiv 7 \pmod{16}$
- $9050775 \geq 85031$
- $\lambda' = 339900 \cdot 2^{560360} = 84975 \cdot 2^{560362}$.

## 9.3 Four-plus-term Hofstadter-like Recurrence

The obvious next step is to consider the four-term generalization of the Hofstadter $Q$-recurrence. The four-term generalization has analogous behavior to the five-term, six-term, seven-term, etc. generalizations. For the remainder of this section, let $G_{d,N}(n)$ denote the $n^{th}$ term in the sequence generated by the recurrence

$$G_d(n) = \sum_{i=1}^{d} G_d(n - G_d(n-i))$$

with the initial condition $\langle 1, 2, 3, \ldots, N \rangle$.

We have the following theorem on weak death.

**Theorem 9.18.** *Let $N$ be a natural number, and let $d \geq 4$. Then, the sequence $G_{d,N}$ weakly dies after $N + \frac{1}{2}d^3 + \frac{3}{2}d + 1$ terms, provided $N \geq d^2 + 3$.*

The proof involves describing all of the terms. Since some of these descriptions are more widely applicable, we extract the following proposition, which consists of the bulk of the proof of Theorem 9.18.

**Proposition 9.19.** *Let $d$ be a positive integer, and let $D = \frac{d^2+d}{2}$. Let $N \geq D$ be another positive integer. Then, we have the following characterization of $G_{d,N}(n)$ for $N - d + 1 \leq n \leq N + (D - d + 1)(d + 1)$:*

1. *For $0 \leq k \leq D - d$, $G_{d,N}(N + k(d+1) + 1) = D + kd$.*

2. *For $-1 \leq k \leq D - d$ and $2 \leq r \leq d + 1$, $G_{d,N}(N + k(d+1) + r) = N + kd + r - 1$, as long as we do not have both $k = D - d$ and $r = d + 1$.*

3. *$G_{d,N}(N + (D - d + 1)(d + 1)) = D + (D - d + 1)d - 1$*

*Proof.* The proof is by induction on the index. We observe that the $k = -1$ possibility in case 2 is correct, as these terms are part of the initial condition.

Now, suppose $N + 1 \leq n \leq N + (D - d + 1)(d + 1) - 1$, and suppose inductively that the formulas hold for all indices less than $n$. There are three cases to consider, which we actually group into two calculations:

$n$ **in case 1:** Suppose $n = N + k(d+1) + 1$ for some $0 \leq k \leq D - d$. We have

$$\begin{aligned}
G_{d,N}(n) &= \sum_{i=1}^{d} G_{d,N}(n - G_{d,N}(n-i)) \\
&= \sum_{i=1}^{d} G_{d,N}(N + k(d+1) + 1 - G_{d,N}(N + k(d+1) + 1 - i)) \\
&= \sum_{i=1}^{d} G_{d,N}(N + k(d+1) + 1 - G_{d,N}(N + (k-1)(d+1) + d + 2 - i)) \\
&= \sum_{r=2}^{d+1} G_{d,N}(N + k(d+1) + 1 - G_{d,N}(N + (k-1)(d+1) + r)) \\
&= \sum_{r=2}^{d+1} G_{d,N}(N + k(d+1) + 1 - (N + (k-1)d + r - 1)) \\
&= \sum_{r=2}^{d+1} G_{d,N}(d + k - r + 2).
\end{aligned}$$

The indices here range from $k+1$ through $k+d$. As long as $k \leq D - d$, all the indices are at most $D$. These all fall in the initial condition, allowing us to conclude

$$\begin{aligned}
G_{d,N}(n) &= \sum_{r=2}^{d+1} G_{d,N}(d + k - r + 2) \\
&= \sum_{r=2}^{d+1} (d + k - r + 2) \\
&= d(d + k + 2) - (D + d) \\
&= d^2 + d + kd - D \\
&= D + kd,
\end{aligned}$$

as required.

$n$ **in case 2 or 3:** Suppose $n = N + k(d+1) + r$ for some $0 \leq k \leq D - d$ and $2 \leq r \leq d + 1$. We have

$$\begin{aligned}
G_{d,N}(n) &= \sum_{i=1}^{d} G_{d,N}(n - G_{d,N}(n-i)) \\
&= \sum_{i=1}^{d} G_{d,N}(N + k(d+1) + r - G_{d,N}(N + k(d+1) + r - i)) \\
&= G_{d,N}(N + k(d+1) + r - G_{d,N}(N + k(d+1) + 1))
\end{aligned}$$

$$
\begin{aligned}
&+\sum_{i=1}^{r-2} G_{d,N}(N+k\,(d+1)+r-G_{d,N}(N+k\,(d+1)+r-i))\\
&+\sum_{i=r}^{d} G_{d,N}(N+k\,(d+1)+r-G_{d,N}(N+k\,(d+1)+r-i))\\
=\;& G_{d,N}(N+k\,(d+1)+r-G_{d,N}(N+k\,(d+1)+1))\\
&+\sum_{i=1}^{r-2} G_{d,N}(N+k\,(d+1)+r-G_{d,N}(N+k\,(d+1)+r-i))\\
&+\sum_{i=r}^{d} G_{d,N}(N+k\,(d+1)+r\\
&-G_{d,N}(N+(k-1)\,(d+1)+r-i+d+1))\\
=\;& G_{d,N}(N+k\,(d+1)+r-(D+kd))\\
&+\sum_{i=1}^{r-2} G_{d,N}(N+k\,(d+1)+r-(N+kd+r-i-1))\\
&+\sum_{i=r}^{d} G_{d,N}(N+k\,(d+1)+r-(N+(k-1)d+r-i+d))\\
=\;& G_{d,N}(N-D+k+r)+\sum_{i=1}^{r-2} G_{d,N}(k+i+1)+\sum_{i=r}^{d} G_{d,N}(k+i).
\end{aligned}
$$

Since $N \geq D$, if $k+r \leq D$, then $G_{d,N}(N-D+k+r) = N-D+k+r$. The only time we would not have $k+r \leq D$ is if $k = D-d$ and $r = d+1$. In this case, we have $G_{d,N}(N-D+k+r) = G_{d,N}(N+1) = D$ (see case 1). The other terms ($k+i+1$ when $i$ is at most $r-2$, which is at most $d-1$, and $k+i$ when $i$ is at most $d$) always fall within the initial condition. So, if $k < D-d$ or $r < d+1$, then we have

$$
\begin{aligned}
G_{d,N}(n) &= G_{d,N}(N-D+k+r)+\sum_{i=1}^{r-2} G_{d,N}(k+i+1)+\sum_{i=r}^{d} G_{d,N}(k+i)\\
&= N-D+k+r+\sum_{i=1}^{r-2}(k+i+1)+\sum_{i=r}^{d}(k+i)\\
&= N-D+k+r+(k+1)\,(r-2)+\frac{(r-2)\,(r-1)}{2}\\
&\quad+(d-r+1)\,k+D-\frac{r\,(r-1)}{2}\\
&= N+dk+r-1,
\end{aligned}
$$

as required. If $k = D-d$ and $r = d+1$, then all of the calculations above go through, except for one term. For that term, instead of $N-D+k+r = N-D+(D-d)+(d+1) =$

$N+1$, we obtain $D$. So, we have (keeping in mind that $k = D-d$ and $r = d+1$)

$$\begin{aligned} G_{d,N}(N+(D-d)(d+1)+d+1) &= G_{d,N}(N+(D-d+1)(d+1)) \\ &= D+(D-d+1)(d-1)+\frac{(d-1)\,d}{2} \\ &\quad + (d-(d+1)+1)(D-d)+D-\frac{(d+1)\,d}{2} \\ &= \frac{1}{2}d^3+\frac{3}{2}d-1 \\ &= D+(D-d+1)\,d-1, \end{aligned}$$

as required. □

The rest of the terms of $G_{d,N}$ are characterized by the following lemma:

**Lemma 9.20.** *Let $d$ be a positive integer. Let $D = \frac{d^2+2}{2}$, and let $D' = (D-d+1)(d+1)$. Let $N \geq D+1$ be a positive integer. For $N+D'+1 \leq n \leq N+D'+d$, we have the following values of $G_{d,N}(n)$:*

1. *$G_{d,N}(N+D'+1) = N+(D-d+1)\,d+1$, provided $d \geq 3$.*
2. *For $2 \leq r \leq d-2$, $G_{d,N}(N+D'+r) = N+(D-d+1)\,d+r-1$, provided $d \geq 3$.*
3. *$G_{d,N}(N+D'+d-1) = D+(D-d+2)\,d-3$, provided $d \geq 3$.*
4. *$G_{d,N}(N+D'+d) = 2N+(D-d+1)(d-1)$, provided $d \geq 4$.*

*Proof.* The proof is by induction on $n$. We prove each case in turn, using Proposition 9.19 whenever applicable. For this reason, we do not actually need a base case. (We also omit straightforward but tedious algebraic manipulations.) So, whenever we examine an $n$ value, we are assuming that this lemma holds for all smaller $n$ values.

$n$ **in case 1:** Here, $n = N+D'+1$. We have

$$\begin{aligned} G_{d,N}(n) &= \sum_{i=1}^{d} G_{d,N}(n-G_{d,N}(n-i)) \\ &= \sum_{i=1}^{d} G_{d,N}(N+D'+1-G_{d,N}(N+D'+1-i)) \\ &= G_{d,N}(N+D'+1-G_{d,N}(N+D')) \end{aligned}$$

$$
\begin{aligned}
&\quad + \sum_{i=2}^{d} G_{d,N}(N + D' + 1 - G_{d,N}(N + D' + 1 - i)) \\
&= G_{d,N}(N + D' + 1 - (D + (D - d + 1)\, d - 1)) \\
&\quad + \sum_{i=2}^{d} G_{d,N}(N + D' + 1 - (N + (D - d)\, d + d - i + 1)) \\
&= G_{d,N}(N - d + 3) + \sum_{i=2}^{d} G_{d,N}(D - d + i + 1).
\end{aligned}
$$

The first term falls in the initial condition as long as $d \geq 3$. The remaining terms all fall in the initial condition, since $N \geq D + 1$. Since this case assumes $d \geq 3$, we can proceed:

$$
\begin{aligned}
G_{d,N}(n) &= G_{d,N}(N - d + 3) + \sum_{i=2}^{d} G_{d,N}(D - d + i + 1) \\
&= N - d + 3 + \sum_{i=2}^{d} (D - d + i + 1) \\
&= N - d + 3 + (d - 1)\,(D - d + 1) + D - 1 \\
&= N + (D - d + 1)\, d + 1,
\end{aligned}
$$

as required.

$n$ **in case 2 or 3:** Here, $n = N + D' + r$ for some integer $2 \leq r \leq d - 1$. We have

$$
\begin{aligned}
G_{d,N}(n) &= \sum_{i=1}^{d} G_{d,N}(n - G_{d,N}(n - i)) \\
&= \sum_{i=1}^{d} G_{d,N}(N + D' + r - G_{d,N}(N + D' + r - i)) \\
&= G_{d,N}(N + D' + r - G_{d,N}(N + D')) \\
&\quad + G_{d,N}(N + D' + r - G_{d,N}(N + D' + 1)) \\
&\quad + \sum_{i=1}^{r-2} G_{d,N}(N + D' + r - G_{d,N}(N + D' + r - i)) \\
&\quad + \sum_{i=r+1}^{d} G_{d,N}(N + D' + r - G_{d,N}(N + D' + r - i)) \\
&= G_{d,N}(N + D' + r - (D + (D - d + 1)\, d - 1)) \\
&\quad + G_{d,N}(N + D' + r - (N + (D - d + 1)\, d + 1))
\end{aligned}
$$

$$
\begin{aligned}
&+\sum_{i=1}^{r-2} G_{d,N}(N+D'+r-(N+(D-d+1)\,d+r-i-1)) \\
&+\sum_{i=r+1}^{d} G_{d,N}(N+D'+r-(N+(D-d)\,d+d+r-i)) \\
=\;& G_{d,N}(N-d+r+2)+G_{d,N}(D-d+r) \\
&+\sum_{i=1}^{r-2} G_{d,N}(D-d+i+2)+\sum_{i=r+1}^{d} G_{d,N}(D-d+i+1).
\end{aligned}
$$

We now consider cases 2 and 3 separetely. If we are in case 2, then $r \leq d-2$. In our calculations, we use earlier versions of this case, and we use case 1. So, we need $d \geq 3$. But, under this assumption, all terms in the last expression above fall in the initial condition. This allows us to proceed:

$$
\begin{aligned}
G_{d,N}(n) &= G_{d,N}(N-d+r+2)+G_{d,N}(D-d+r) \\
&\quad+\sum_{i=1}^{r-2} G_{d,N}(D-d+i+2)+\sum_{i=r+1}^{d} G_{d,N}(D-d+i+1) \\
&= (N-d+r+2)+(D-d+r) \\
&\quad+\sum_{i=1}^{r-2} (D-d+i+2)+\sum_{i=r+1}^{d} (D-d+i+1) \\
&= (N-d+r+2)+(D-d+r)+(r-2)\,(D-d+2) \\
&\quad+\frac{(r-2)\,(r-1)}{2}+(d-r)\,(D-d+1)+D-\frac{r\,(r+1)}{2} \\
&= N+(D-d+1)\,d+r-1,
\end{aligned}
$$

as required.

If we are in case 3, then $r = d-1$. In this case, $N-d+r+2 = N+1$, so $G_{d,N}(N-d+r+2) = D$. By the same logic we use for case 2, we need $d \geq 3$. But, under this condition, all the other terms are in the initial condition. So, we proceed:

$$
\begin{aligned}
G_{d,N}(n) &= G_{d,N}(N-d+(d-1)+2)+G_{d,N}(D-d+(d-1)) \\
&\quad+\sum_{i=1}^{d-3} G_{d,N}(D-d+i+2)+\sum_{i=d}^{d} G_{d,N}(D-d+i+1)
\end{aligned}
$$

$$\begin{aligned}
&= D + (D - d + (d-1)) + \sum_{i=1}^{d-3} (D - d + i + 2) + (D - d + d + 1) \\
&= D + (D-1) + (d-3)(D - d + 2) + \frac{(d-3)(d-2)}{2} + (D+1) \\
&= D + (D - d + 2)\, d - 3,
\end{aligned}$$

as required.

$n$ **in case 4:** Here, $n = N + D' + d$. We have

$$\begin{aligned}
G_{d,N}(n) &= \sum_{i=1}^{d} G_{d,N}(n - G_{d,N}(n-i)) \\
&= \sum_{i=1}^{d} G_{d,N}(N + D' + d - G_{d,N}(N + D' + d - i)) \\
&= G_{d,N}(N + D' + d - G_{d,N}(N + D')) \\
&\quad + G_{d,N}(N + D' + d - G_{d,N}(N + D' + 1)) \\
&\quad + G_{d,N}(N + D' + d - G_{d,N}(N + D' + d - 1)) \\
&\quad + \sum_{r=2}^{d-2} G_{d,N}(N + D' + d - G_{d,N}(N + D' + r)) \\
&= G_{d,N}(N + D' + d - (D + (D - d + 1)\, d - 1)) \\
&\quad + G_{d,N}(N + D' + d - (N + (D - d + 1)\, d + 1)) \\
&\quad + G_{d,N}(N + D' + d - (D + (D - d + 2)\, d - 3)) \\
&\quad + \sum_{r=2}^{d-2} G_{d,N}(N + D' + d - (N + (D - d + 1)\, d + r - 1)) \\
&= G_{d,N}(N+2) + G_{d,N}(D) + G_{d,N}(N - d + 4) + \sum_{r=2}^{d-2} G_{d,N}(D - r + 2).
\end{aligned}$$

If $d \geq 2$, then $G_{d,N}(N+2) = N+1$. Also, we know that $D \leq N$, so $G_{d,N}(D) = D$. Similarly, the terms in the summation all fall in the initial condition, as the indices are all at most $D$. If $d \geq 4$, the remaining term also falls in the initial condition. So, if $d \geq 4$, we can proceed:

$$\begin{aligned}
G_{d,N}(n) &= G_{d,N}(N+2) + G_{d,N}(D) + G_{d,N}(N - d + 4) + \sum_{r=2}^{d-2} G_{d,N}(D - r + 2) \\
&= N + 1 + D + N - d + 4 + \sum_{r=2}^{d-2} (D - r + 2)
\end{aligned}$$

$$= 2N + D - d + 5 + (d-3)(D+2) - \frac{(d-2)(d-1)}{2} + 1$$
$$= 2N + (D - d + 1)(d-1),$$

as required.

□

We now complete the proof of Theorem 9.18.

*Proof.* Let $d \geq 4$ and $N \geq d^2 + 3$ be integers. Let $D$ and $D'$ be as in Lemma 9.20. Proposition 9.19 and Lemma 9.20 (and the initial condition) give the values of the first $N + D' + d$ terms of $G_{d,N}$. (This is valid as the most restrictive conditions in any of these cases are $d \geq 4$ and $N \geq D + 1$, which are satisfied here.) So, the sequence lives at least this long. Note that $D' = \frac{1}{2}d^3 + \frac{1}{2}d + 1$. So, $N + D' + d = N + \frac{1}{2}d^3 + \frac{3}{2}d + 1$. This is exactly the number of terms we wish to show that $G_{d,N}$ lives for, so all that remains is to show that $G_{d,N}$ weakly dies at this point.

According to Lemma 9.20, $G_{d,N}(N + D' + d) = 2N + (D - d + 1)(d - 1)$. This has a $2N$ in it, so for sufficiently large $N$, it will exceed $N + D' + d + 1$. If this happens, the sequence weakly dies. This inequality is satisfied precisely when $N \geq d^2 + 3$, which is what we have here. So, our sequence weakly dies at index $N + \frac{1}{2}d^3 + \frac{3}{2}d + 1$, as required. □

In Sections 9.1 and 9.2, we also have a characterization of when the sequences considered there strongly die. Such a characterization for $G_{d,N}$ when $d \geq 4$ seems much more difficult. Empirically, under the strong death convention, these sequences frequently exhibit quasilinear behavior for awhile. But, there is a point at which such behavior ends, after which the terms look completely chaotic. See Figures 9.7 through 9.10 for some examples. These figures emphasize that a small change in the initial conditions can lead to a large change in the behavior, despite some global similarities.

For some values of $d$ (including $d \in \{4, 5, 6, 8, 12, 13\}$), the sequence $G_{d,N}$ settles into a long quasilinear stretch with period $d + 1$. All but two of the interleaved sequences have slope $\frac{d}{d+1}$. One of the other two has slope 1, and the other has slope $\frac{d-1}{d+1}$. For other

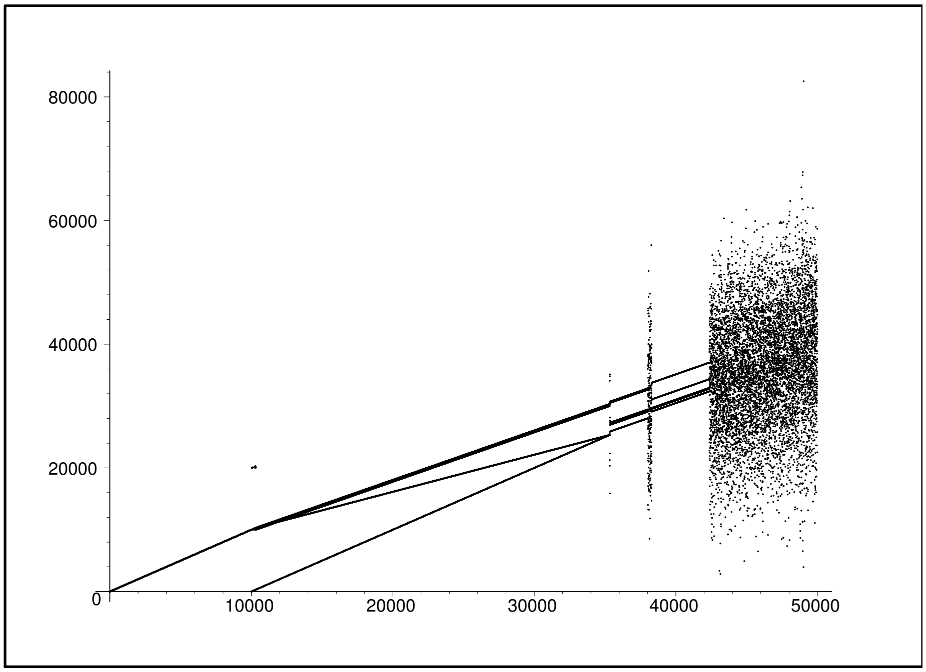


Figure 9.7: The first 50000 terms of $G_{4,10000}$ (A283889)

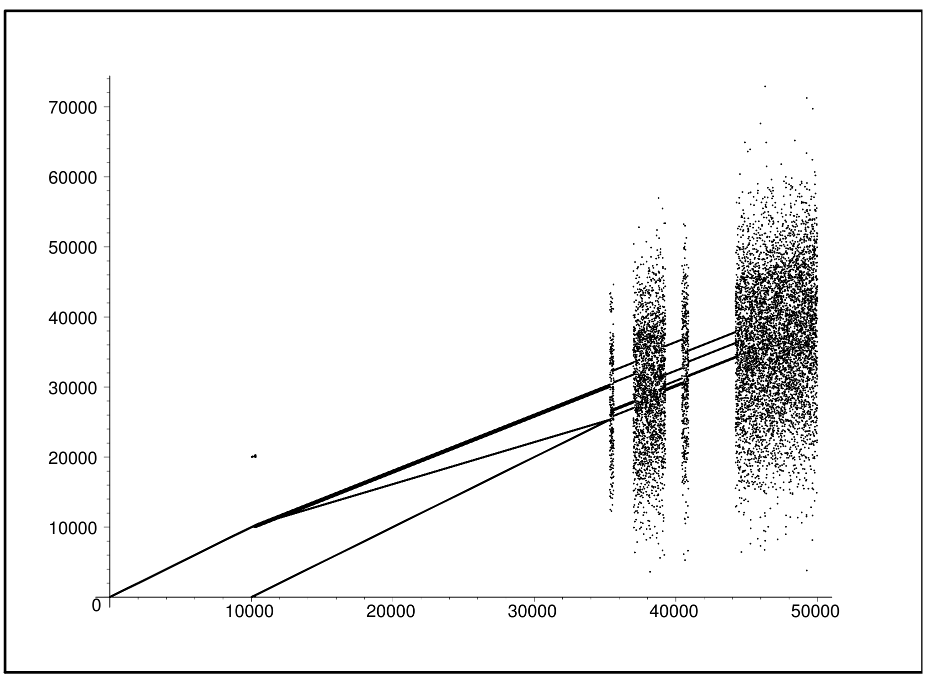


Figure 9.8: The first 50000 terms of $G_{4,10001}$ (A283890)

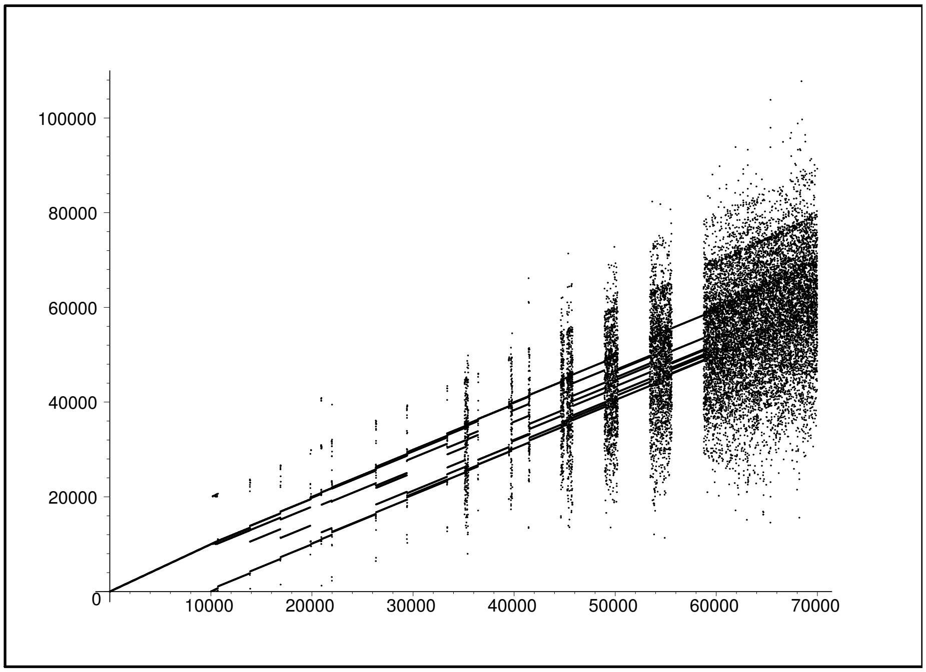


Figure 9.9: The first 70000 terms of $G_{7,10000}$ (A283891)

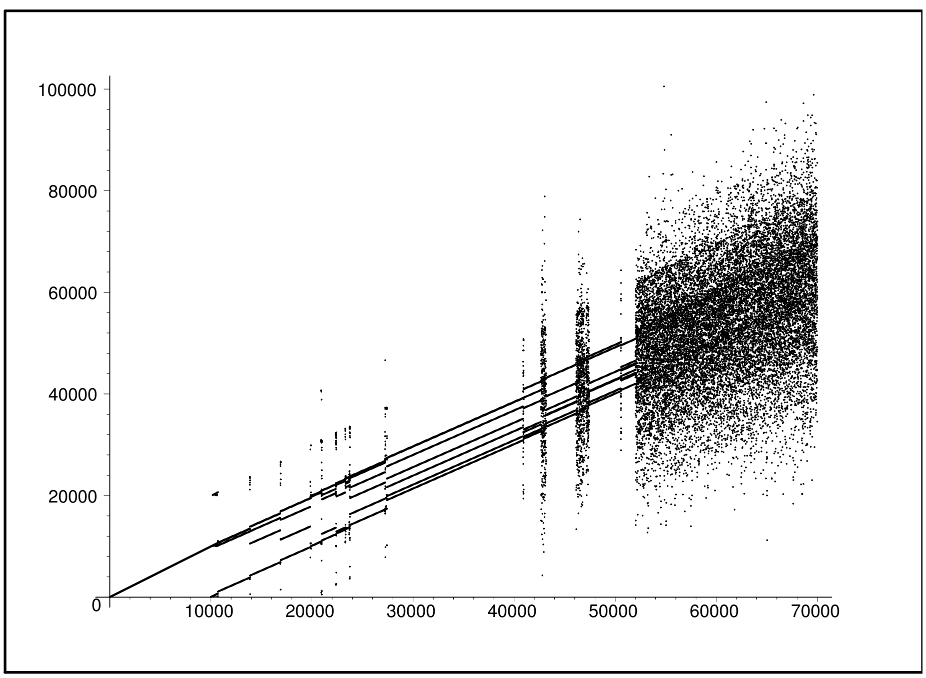


Figure 9.10: The first 70000 terms of $G_{7,10001}$ (A283892)

values of $d$ (including $d \in \{7, 9, 10, 11, 14, 15, 16\}$), the sequence reaches a comparatively short quasilinear stretch, also with period $d+1$. But, all of the component sequences here have slope $\frac{d}{d+1}$. Following the first quasilinear stretch, both types of initial stretch sometimes lead to multiple additional stretches of this second type, interrupted by chaotic interludes. (See Figure 9.7 for a good example of this with $d = 4$.) Eventually, though, the quasilinear parts appear to end and devolve in to complete chaos. When employing our symbolic method, eventual chaos manifests itself as a rapid dependence on stronger and stronger congruence constraints on $N$ (for example, each successive term depending on $N$ mod the next power of 2).

## 9.4 Failure of a Slow Solution to Generalize

In Chapter 6, we analyze the recurrence $B(n) = B(n - B(n-1)) + B(n - B(n-2)) + B(n - B(n-3))$ with the initial condition $\langle 1, 2, 3, 4, 5\rangle$. We discover there that the resulting sequence is slow. (See Section 2.3 in Chapter 2 for an introduction to slow sequences.) In this section, we discuss possible generalizations of this result. In particular, we use results from Section 9.3 to show that obvious generalizations of our solution do not generate a slow sequence.

According to the work of Isgur et al. [20], our $B$-sequence is the fundamental member of an infinite family of slow sequences with similar recurrences. (The next one satisfies the recurrence $B'(n) = B'(n - B'(n-2)) + B'(n - B'(n-4)) + B'(n - B'(n-6))$.) As mentioned in Chapter 2, this family and the family resulting from the $V$-sequence comprise the only known examples of slow Hofstadter-like sequences with all recurrence terms of the form $D(n - D(n-i))$ for some $i$. We have conducted a search for other such sequences without finding another (nontrivial) example. (See `http://github.com/nhf216/thesis/slowsearch.txt` for code to conduct such a search, and see `http://github.com/nhf216/thesis/slowseqs.txt` for some slow solutions found.) An obvious idea would be to generalize the $B$-recurrence to the $d$-term recurrence

$$G_d(n) = \sum_{i=1}^{d} G_d(n - G_d(n-i))$$

that appears in Section 9.3. If $d = 1$, the initial condition $\langle 1 \rangle$ generates the all-ones sequence, which, while technically slow, is not particularly interesting. We have the following non-existence result:

**Theorem 9.21.** *The $B$-sequence is the only nontrivial slow sequence resulting from a recurrence $G_d$ with an initial condition of the form $\langle 1, 2, 3, \ldots, N \rangle$ for some $N$.*

The bulk of the Theorem 9.21 follows from the following proposition:

**Proposition 9.22.** *Suppose $d \geq 4$. Let $D = \frac{d^2+d}{2}$. The sequence $G_{d,D}$ (see Section 9.3 for a definition of this notation) satisfies*

$$G_{d,D}\left(\frac{1}{2}d^3 + \frac{1}{2}d^2 + 2d + 1\right) = G_{d,D}\left(\frac{1}{2}d^3 + \frac{1}{2}d^2 + 2d\right) + 2.$$

*In particular, the sequence has a jump of difference 2, so it is not slow. Moreover, the sequence is slow before this jump.*

*Proof.* Proposition 9.19 in the previous section characterizes the sequence $G_{d,D}$ through $G_{d,D}(D + (D - d + 1)(d + 1))$. Since $N = D$, it is easy to see that the successive differences are all 0 or 1 throughout these terms.

Now, for simplicity of notation, let $D' = (D - d + 1)(d + 1)$. We will now show that, for $1 \leq r \leq d - 2$, $G_{d,D}(D + D' + r) = D + (D - d + 1)d + r - 1$. Inductively, suppose this holds for all $r' < r$. We now calculate

$$\begin{aligned}
&G_{d,D}(D + D' + r) \\
&= \sum_{i=1}^{d} G_{d,D}(D + D' + r - G_{d,D}(D + D' + r - i)) \\
&= \sum_{i=1}^{r} G_{d,D}(D + D' + r - G_{d,D}(D + D' + r - i)) \\
&\quad + \sum_{i=r+1}^{d} G_{d,D}(D + D' + r - G_{d,D}(D + D' + r - i)) \\
&= \sum_{i=1}^{r} G_{d,D}(D + D' + r - (D + (D - d + 1)d + r - i - 1)) \\
&\quad + \sum_{i=r+1}^{d} G_{d,D}(D + D' + r - (D + (D - d)d + (d + 1 + r - i) - 1))
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{r} G_{d,D}(D' + i + 1 - (D - d + 1)\, d) + \sum_{i=r+1}^{d} G_{d,D}(D' + i - (D - d + 1)\, d) \\
&= \sum_{i=1}^{r} G_{d,D}(D - d + i + 2) + \sum_{i=r+1}^{d} G_{d,D}(D - d + i + 1) \\
&= D + \sum_{i=1}^{r} (D - d + i + 2) + \sum_{i=r+1}^{d-1} (D - d + i + 1) \\
&= Dd - (d-1)\, d + r + (d-1) + \frac{d^2 - d}{2} \\
&= (D - d + 1)\, d + r - 1 + \left( d + \frac{d^2 - d}{2} \right) \\
&= D + (D - d + 1)\, d + r - 1,
\end{aligned}
$$

as required. The above calculation is also valid for $r = d - 1$, except that $G_{d,D}(D - d + i + 2)$ would be $G_{d,D}(D+1)$ when $i = d - 1$. Recall that $G_{d,D}(D+1) = D$, rather than $D + 1$. So, we obtain $G_{d,D}(D + D' + d - 1) = D + (D - d + 1)d + d - 3$.

We now compute

$$
\begin{aligned}
&G_{d,D}(D + D' + d) \\
&= \sum_{i=1}^{d} G_{d,D}(D + D' + d - G_{d,D}(D + D' + d - i)) \\
&= G_{d,D}(D + D' + d - G_{d,D}(D + D' + d - 1)) \\
&\quad + \sum_{i=2}^{d} G_{d,D}(D + D' + d - G_{d,D}(D + D' + d - i)) \\
&= G_{d,D}(D + D' + d - (D + (D - d + 1)\, d + d - 3)) \\
&\quad + \sum_{i=2}^{d} G_{d,D}(D + D' + d - (D + (D - d + 1)\, d + d - i - 1)) \\
&= G_{d,D}(D' - (D - d + 1)\, d + 3) + \sum_{i=2}^{d} G_{d,D}(D' - (D - d + 1)\, d + i + 1) \\
&= G_{d,D}(D - d + 4) + \sum_{i=2}^{d} G_{d,D}(D - d + i + 2) \\
&= G_{d,D}(D - d + 4) + \sum_{i=2}^{d-2} G_{d,D}(D - d + i + 2) + G_{d,D}(D+1) + G_{d,D}(D+2) \\
&= D - d + 4 + \sum_{i=2}^{d-2} (D - d + i + 2) + 2D + 1
\end{aligned}
$$

$$
\begin{aligned}
&= 3D - d + 5 + (d-3)\,D - d(d-3) + 2\,(d-3) + \left(\frac{(d-2)\,(d-1)}{2} - 1\right) \\
&= Dd - d + 4 - d\,(d-3) + 2\,(d-3) + \frac{(d-2)\,(d-1)}{2} \\
&= Dd - d + 4 - d\,(d-1) + 2d + 2d - 6 + (D - d - (d-1)) \\
&= D + (D - d + 1)\,d + d - 1.
\end{aligned}
$$

(Observe that these calculations are only valid because $d \geq 4$, as otherwise $D - d + 4$ would be larger than $D$.) So, we have $G_{d,D}(D + D' + d) = G_{d,D}(D + D' + d - 1) + 2$. Recalling the values of $D$ and $D'$, we have that $D + D' + d = \frac{1}{2}d^3 + \frac{1}{2}d^2 + 2d + 1$, as required. □

We will now complete the proof of Theorem 9.21.

*Proof.* Fix a positive integer $N$. Consider the sequence $G_{d,N}$. Suppose that the sequence we obtain is slow. Clearly, we need $N \geq d$, or else $G_{d,N}$ is undefined. Supposing that $N \geq d$, we have

$$
\begin{aligned}
G_{d,N}(N+1) &= \sum_{i=1}^{d} G_{d,N}(N + 1 - G_{d,N}(N + 1 - i)) \\
&= \sum_{i=1}^{d} G_{d,N}(N + 1 - (N + 1 - i)) \\
&= \sum_{i=1}^{d} G_{d,N}(i) \\
&= \sum_{i=1}^{d} i \\
&= \frac{d^2 + d}{2}.
\end{aligned}
$$

So, unless $N \in \left\{\frac{d^2+d}{2} - 1, \frac{d^2+d}{2}\right\}$, we would not have $G_{d,N}(N+1) - G_{d,N}(N) \in \{0, 1\}$. According to Proposition 9.22, $N = \frac{d^2+d}{2}$ does not result in a slow sequence for $d \geq 4$. Similarly, $N = \frac{d^2+d}{2} - 1$ does not result in a slow sequence for $k \geq 4$, as this sequence is identical to the one for $N = \frac{d^2+d}{2}$ (since the first $N$ terms are the same). So, we must have $d \leq 3$. The case $d = 1$ results in a trivial sequence, $d = 2$ give the Hofstadter $Q$-sequence (which is not slow), and $d = 3$ gives our $B$-sequence. Therefore, the $B$-sequence is the only nontrivial slow sequence of the form $G_{d,N}$, as required. □

# Chapter 10

# Nested Recurrences with other Symbolic Initial Conditions

Chapter 9 studies initial conditions of the form $\langle 1, 2, 3, \ldots, N\rangle$. There are a multitude of other types of initial conditions to study. In this chapter, we consider the Hofstadter $Q$-recurrence with initial conditions of the form $\langle N, 2\rangle$, $\langle 2, N\rangle$, $\langle N, 4, N, 4\rangle$, and $\langle 4, N, 4, N\rangle$. The motivations for studying these initial conditions are as follows:

- Each of these is a parametrized family of *constant-length* initial conditions.
- The terms immediately after the initial condition continue alternating between 2 or 4 and $N$, through index $N$. In this way, the initial condition generates a long, temporary period-2 quasilinear solution to the $Q$-recurrence.

This second motivation explains why we do not consider alternations of threes and $N$'s.

## 10.1 Twos and $N$'s

We have the following result for the initial condition $\langle N, 2\rangle$:

**Theorem 10.1.** *Let $N$ be a natural number. Let $Q_{N,2}$ be the sequence resulting from the Hofstadter $Q$-recurrence and the initial condition $\langle N, 2\rangle$. We have the following cases:*

1. *If $N = 1$, then $Q_{N,2}$ is the Hofstadter $Q$-sequence shifted by $1$ term.*
2. *If $N$ is even, then we have a quasilinear solution with period $N$. Each period consists of $\frac{N}{2}$ repeated blocks of of length $2$. In the description of these solutions that follows, $0 \leq \ell < \frac{N}{2}$.*

- $Q_{N,2}(Nk + 2\ell) = 2$
- $Q_{N,2}(Nk + 2\ell + 1) = N(k+1)$.

3. *If $N \geq 25$ and $N \equiv 3 \pmod 4$, then $Q_{N,2}$ strongly dies after $5N + 11$ terms.*

4. *If $N \geq 75$ and $N \equiv 5 \pmod{12}$, then $Q_{N,2}$ strongly dies after $28N + 64$ terms.*

5. *If $N \geq 51$ and $N \equiv 1, 9, 13, 21 \pmod{24}$, then $Q_{N,2}$ persists forever and is eventually quasilinear with period $24N + 34$. The pattern begins at index $53N + 107$.*

*Proof.* In the cases other than case 5 where the sequence does not strongly die, the result is easy to prove by induction. In the cases where the sequence dies, the code in `http://github.com/nhf216/thesis/nonstdhof.txt` can be used to generate closed forms for all of the terms in the sequence. A zero occurs in the indicated position, which causes the sequence to strongly die.

Case 5 is much more complicated. Exploring these sequences with `nonstdhof.txt` does not terminate. But, eventually, provided $N \geq 51$, an apparent "second-order" pattern develops among the blocks explored (see p. 97 for a discussion of higher-order patterns). The second-order pattern here is described by the following:
Let $K_0 = 72N+99$, $K_1 = 73N+99$, $K_2 = 74N+103$, $K_3 = 75N+101$, $K_4 = 76N+103$, and $K_5 = 77N + 103$. Recursively, let $K_i = K_{i-6} + \lambda$ for $i \geq 6$ . The sequence $Q_{N,2}$, beginning at index $K_0$, consists of the following structure for each $i \geq 0$:

- A pattern $Q_{N,2}(K_{6i} + 2k) = K_{6i+1} - 1$, $Q_{N,2}(K_{6i} + 2k + 1) = 2$ persisting through index $K_{6i+1} - 1$.
- Five sporadic terms, followed by a pattern $Q_{N,2}(K_{6i+1} + 2k) = 2N + 4$, $Q_{N,2}(K_{6i+1} + 2k + 1) = K_{6i} + 1$ persisting through index $K_{6i+2} - 1$.
- Four sporadic terms, followed by a pattern $Q_{N,2}(K_{6i+2} + 2k) = K_{6i+3} - 1$, $Q_{N,2}(K_{6i+2} + 2k + 1) = 2$ persisting through index $K_{6i+3} - 1$.
- Twenty-one sporadic terms, followed by a pattern $Q_{N,2}(K_{6i+3} + 2k) = 4N + 4$, $Q_{N,2}(K_{6i+3} + 2k + 1) = K_{6i} + 1$ persisting through index $K_{6i+4} - 1$.

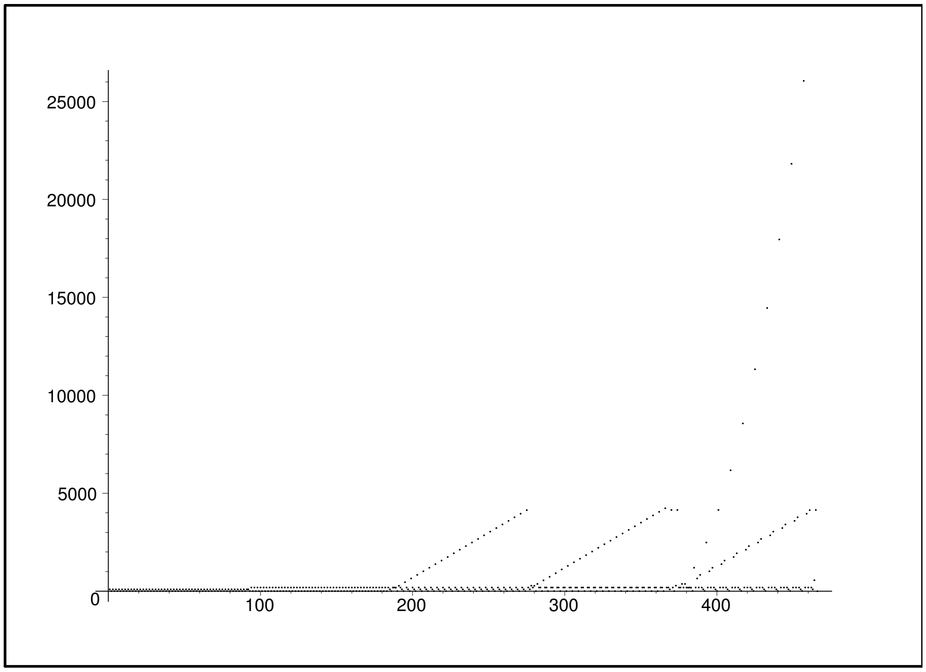


Figure 10.1: All 466 terms of $Q_{91,2}$ (A283897)

- Forty sporadic terms, followed by a pattern $Q_{N,2}(K_{6i+4} + 2k) = K_{6i+5} - 1$, $Q_{N,2}(K_{6i+4} + 2k + 1) = 2$ persisting through index $K_{6i+5} - 1$.

- Thirty-nine sporadic terms, followed by a pattern $Q_{N,2}(K_{6i+5} + 2k) = 2$, $Q_{N,2}(K_{6i+5} + 2k + 1) = K_{6(i+1)} - 1$ persisting through index $K_{6(i+1)} - 1$.

The procedure `ExploreAllN2` in `http://github.com/nhf216/thesisN4N4_explore.txt` carries out the inductive proof of this structure. Given these patterns and the preceding terms, it can be seen that the quasilinear pattern extends all the way back to index $53N+107$, as required. See files `http://github.com/nhf216/thesisN2N2_mod24_1.txt`, `http://github.com/nhf216/thesisN2N2_mod24_9.txt`, `http://github.com/nhf216/thesisN2N2_mod24_13.txt`, and `http://github.com/nhf216/thesisN2N2_mod24_21.txt` for complete descriptions of the sequences in all four possibilities in case 5. □

See Figures 10.1, 10.2, and 10.3 for plots of $Q_{91,2}$, $Q_{89,2}$, and $Q_{57,2}$ which are covered by the last three cases of Theorem 10.1.

For the other odd values of $N$, most $Q_{N,2}$ sequences strongly die. If $N = 3$, then

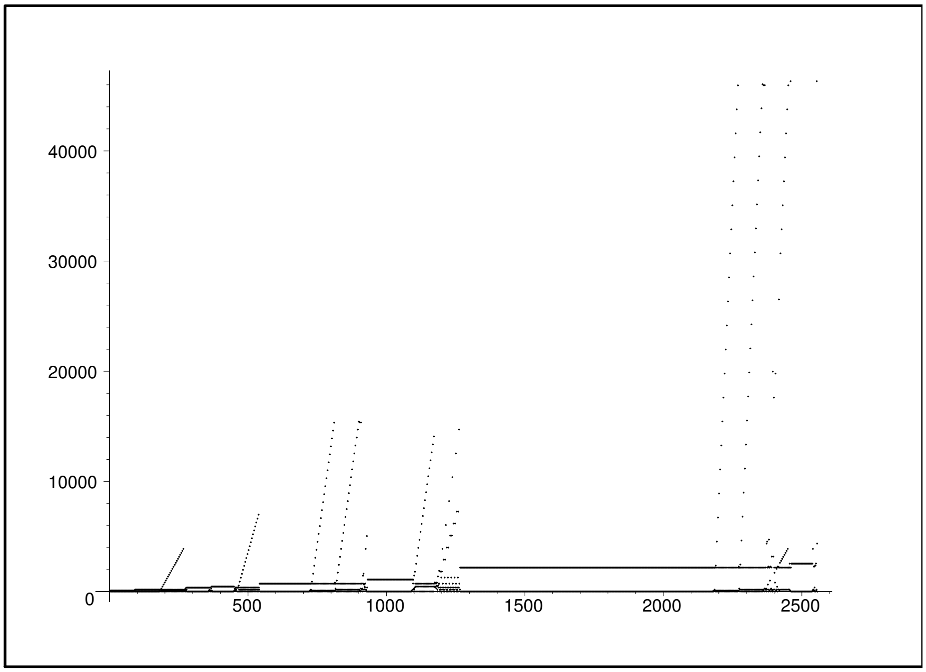


Figure 10.2: All 2556 terms of $Q_{89,2}$ (A283896)

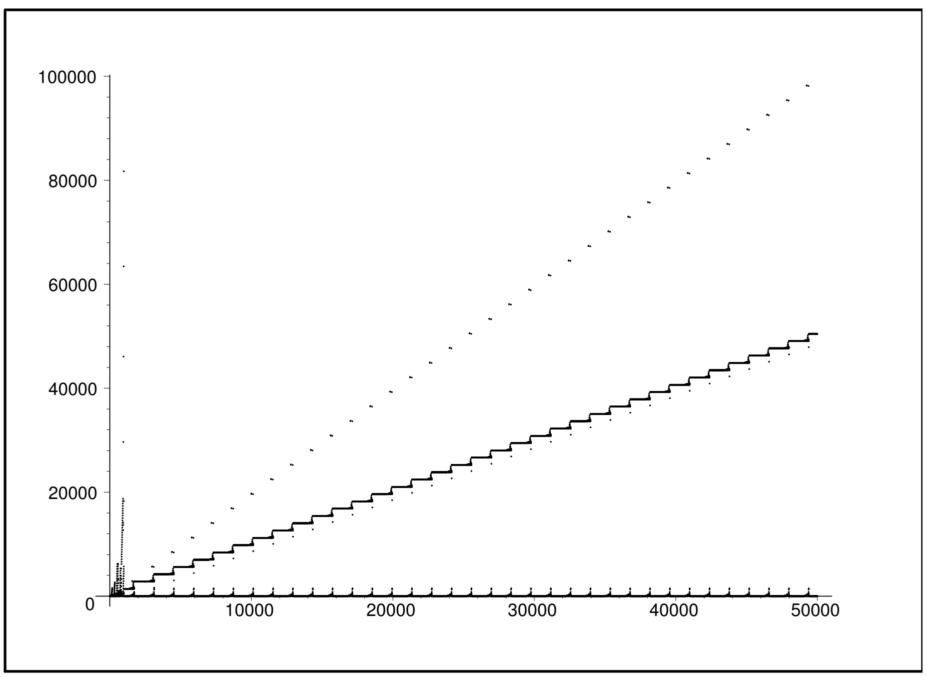


Figure 10.3: First 50000 terms of $Q_{57,2}$ (A278068)

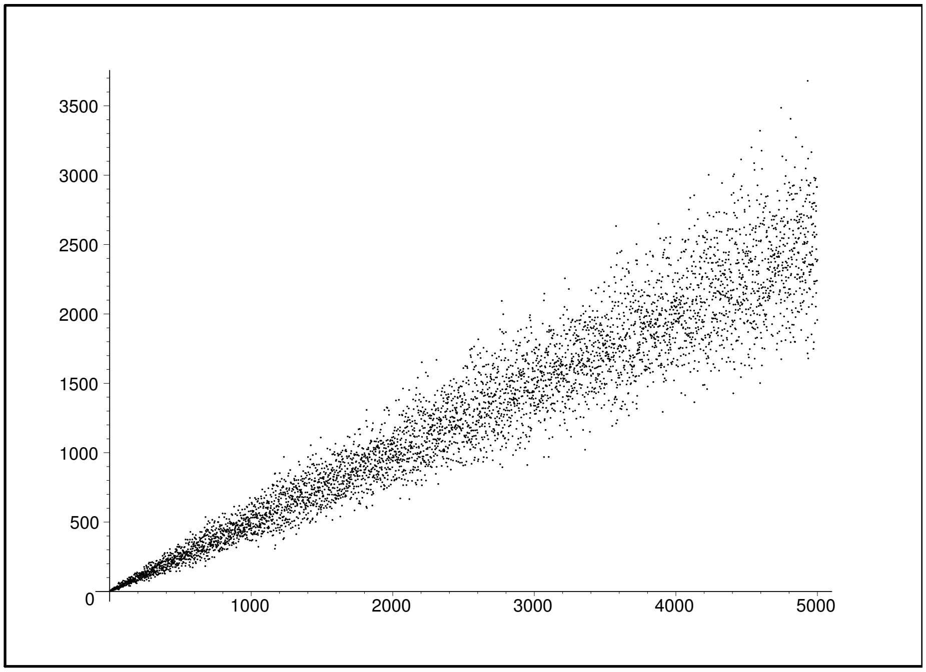


Figure 10.4: First 5000 terms of $Q_{3,2}$ (A283893)

the sequence looks chaotic (Figure 10.4). If $N \in \{5, 17, 41\}$, the sequence $Q_{N,2}$ lives forever, and we can actually predict all of the terms. These results closely resemble the analyses we had for $B_{193}$ and $B_{3442}$ in Subsection 9.2.2 in Chapter 9. The key observation that we exploit is that, for any $\lambda > K \geq 0$, an initial condition of the form $\langle a_1, a_2, a_3, \ldots, a_{K-1}, a_K, \lambda, 2\rangle$ results in alternating $\lambda$'s with twos through index $\lambda$. (This observation is analogous to Proposition 9.10, though it is much simpler.) We call such a pattern a $\lambda$-alternation.

The $N = 5$ case is handled by the following proposition:

**Proposition 10.2.** *Let $K_0 = 25$, and recursively let $K_i = 2K_{i-1} + 1$ for $i \geq 1$. The sequence $Q_{5,2}$, beginning at index 26, consists almost entirely of $(2K_i - 17)$-alternations. The $i^{th}$ alternation persists until index $2K_i - 17 = K_{i+1} - 18$, after which there are 18 sporadic terms that are parametrized by $i$.*

*Proof.* For $K \geq 25$, let $Q_{5,2,K}$ denote the sequence generated by the $Q$-recurrence with initial condition

$$\langle Q_{5,2}(1), Q_{5,2}(2), \ldots, Q_{5,2}(24), Q_{5,2}(25), a_{26}, a_{27}, \ldots, a_{K-1}, a_K, 2K - 17, 2\rangle,$$

where $a_{26}$ through $a_K$ are arbitrary integers. We claim the following about $Q_{5,2,K}$:

1. A $(2K-17)$-alternation persists from index $K+1$ through index $2K-17$.

2. This sequence is an instance of $Q_{5,2,2K+1}$. (By *instance*, we mean that some of the $a_j$'s have been assigned specific values.)

Item 1 is clear. To prove item 2, we use the procedure `KExplore5` in `http://github.com/nhf216/thesis/N4N4_explore.txt` to generate terms and find and prove the next pattern that $Q_{5,2,K}$ reaches. Provided $K \geq 22$, this pattern is precisely a $(4K-15)$-alternation beginning at index $2K+2$. Since the form of the general sequence $Q_{5,2,K}$ is completely determined by the first 25 and final two terms of the initial condition, it is clear that $Q_{5,2,K}$ is an instance of $Q_{5,2,2K+1}$, as required.

Since $Q_{5,2}$ itself is $Q_{5,2,K_0}$ (and $K_0 \geq 22$), Proposition 10.2 follows. $\square$

The 18 sporadic terms between each alternation of $Q_{5,2}$ are as follows:

- $Q_{5,2}(2K_i-16) = 2K_i-12$
- $Q_{5,2}(2K_i-15) = 2$
- $Q_{5,2}(2K_i-14) = 2K_i-12$
- $Q_{5,2}(2K_i-13) = 2$
- $Q_{5,2}(2K_i-12) = 2K_i-12$
- $Q_{5,2}(2K_i-11) = 7$
- $Q_{5,2}(2K_i-10) = 4$
- $Q_{5,2}(2K_i-9) = 2K_i-10$
- $Q_{5,2}(2K_i-8) = 2K_i-10$
- $Q_{5,2}(2K_i-7) = 10$
- $Q_{5,2}(2K_i-6) = 2K_i-10$
- $Q_{5,2}(2K_i-5) = 7$
- $Q_{5,2}(2K_i-4) = 14$
- $Q_{5,2}(2K_i-3) = 6$
- $Q_{5,2}(2K_i-2) = 4K_i-22$
- $Q_{5,2}(2K_i-1) = 10$
- $Q_{5,2}(2K_i) = 4$
- $Q_{5,2}(2K_i+1) = 2K_i-4$

For a plot of $Q_{5,2}$, see Figure 10.5.

The $N = 17$ case is handled by the following proposition:

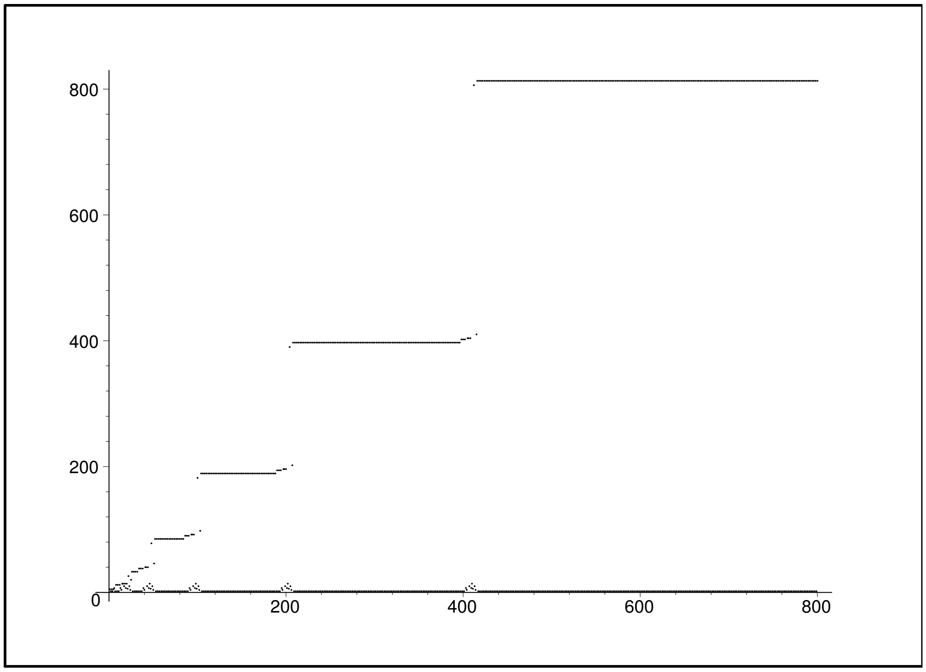


Figure 10.5: First 800 terms of $Q_{5,2}$ (A278066)

**Proposition 10.3.** *Let* $K_0 = 694$*, and recursively let* $K_i = 2K_{i-1} - 100$ *for* $i \geq 1$*. The sequence* $Q_{17,2}$*, beginning at index* $695$*, consists almost entirely of* $(2K_i - 202)$*-alternations. The* $i^{th}$ *alternation persists until index* $2K_i - 202 = K_{i+1} - 102$*, after which there are* $102$ *sporadic terms that are parametrized by* $i$*.*

*Proof.* For $K \geq 694$, let $Q_{17,2,K}$ denote the sequence generated by the $Q$-recurrence with initial condition

$$\langle Q_{17,2}(1), Q_{17,2}(2), \ldots, Q_{17,2}(693), Q_{17,2}(694), a_{695}, a_{696}, \ldots, a_{K-1}, a_K, 2K - 202, 2\rangle,$$

where $a_{695}$ through $a_K$ are arbitrary integers. We claim the following about $Q_{17,2,K}$:

1. A $(2K - 202)$-alternation persists from index $K + 1$ through index $2K - 202$.

2. This sequence is an instance of $Q_{17,2,2K+1}$.

Item 1 is clear. To prove item 2, we use the procedure `KExplore17` in `http://github.com/nhf216/thesis/N4N4_explore.txt` to generate terms and find and prove the next pattern that $Q_{17,2,K}$ reaches. Provided $K \geq 236$, this pattern is precisely a $(4K - 402)$-alternation beginning at index $2K - 100$. Since the form of the general

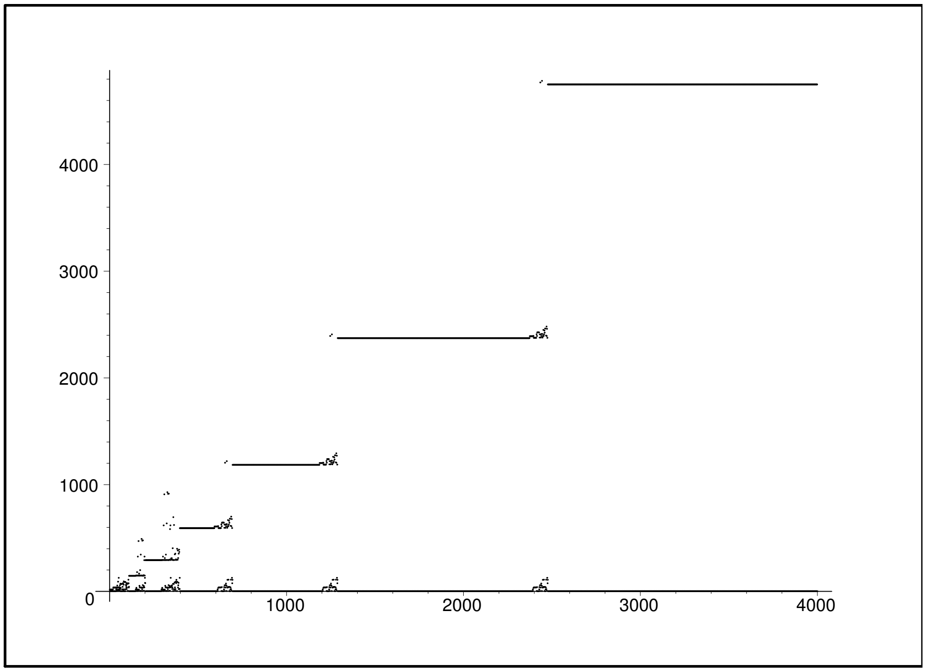


Figure 10.6: First 4000 terms of $Q_{17,2}$ (A283894)

sequence $Q_{17,2,K}$ is completely determined by the first 694 and final two terms of the initial condition, it is clear that $Q_{17,2,K}$ is an instance of $Q_{17,2,2K-100}$, as required.

Since $Q_{17,2}$ itself is $Q_{17,2,K_0}$ (and $K_0 \geq 236$), Proposition 10.3 follows. □

For a listing of the 102 sporadic terms, see `http://github.com/nhf216/thesis/N2N2_17_2_Sporadic.txt`. Interestingly, these terms contain within them some other short alternations, and there are some alternations before index 694 that "almost" follow the pattern. For a plot of $Q_{17,2}$, see Figure 10.6.

The $N = 41$ case is somewhat more complicated. It is handled by the following proposition:

**Proposition 10.4.** *Let* $K_0 = 2639$ *and* $\lambda_0 = 4930$*. Recursively, for* $i \geq 1$*, define* $K_i$ *and* $\lambda_i$ *as follows:*

- *If* $\lambda_{i-1} - K_{i-1} \equiv 0 \pmod 2$*, let* $K_i = \lambda_{i-1} + 562$ *and* $\lambda_i = 2\lambda_i + 41$*.*
- *If* $\lambda_{i-1} - K_{i-1} \equiv 1 \pmod 2$*, let* $K_i = \lambda_{i-1} + 224$ *and* $\lambda_i = 2\lambda_i + 8$*.*

*The sequence* $Q_{41,2}$*, beginning at index* 2640*, consists almost entirely of* $\lambda_i$*-alternations.*

*The $i^{th}$ alternation persists until index $\lambda_i$, after which there are 562 or 224 sporadic terms (depending on the aforementioned parities) that are parametrized by $i$.*

*Proof.* For $K \geq 2639$ and $\lambda \geq K+2$, let $Q_{41,2,K,\lambda}$ denote the sequence generated by the $Q$-recurrence with initial condition

$$\langle Q_{41,2}(1), Q_{41,2}(2), \ldots, Q_{41,2}(2638), Q_{41,2}(2639), a_{2640}, a_{2641}, \ldots, a_{K-1}, a_K, \lambda, 2\rangle,$$

where $a_{2640}$ through $a_K$ are arbitrary integers. We claim the following about $Q_{41,2,K,\lambda}$:

1. A $\lambda$-alternation persists from index $K+1$ through index $\lambda$.
2. If $\lambda - K$ is even, then this sequence is an instance of $Q_{41,2,\lambda+562,2\lambda+41}$. If $\lambda - K$ is odd, then this sequence is an instance of $Q_{41,2,\lambda+224,2\lambda+8}$.

Item 1 is clear. To prove item 2, we use the procedure `KExplore41` in `http://github.com/nhf216/thesis/N4N4_explore.txt` to generate terms and find and prove the next pattern that $Q_{41,2,K,\lambda}$ reaches. Provided $\lambda \geq 3082$, this pattern is precisely the appropriate one of a $(\lambda+8)$ or $(\lambda+41)$-alternation beginning at the appropriate index of $\lambda+562$ or $\lambda+224$. (We actually only need $\lambda \geq 675$ in the odd difference case.) Since the form of the general sequence $Q_{41,2,K,\lambda}$ is completely determined by the first 2639 and final two terms of the initial condition, it is clear that $Q_{41,2,K,\lambda}$ is an instance of $Q_{41,2,\lambda+562,2\lambda+41}$ or $Q_{41,2,\lambda+224,2\lambda+8}$, as required.

Since $Q_{41,2}$ itself is $Q_{41,2,K_0,\lambda_0}$ (and $\lambda_0 \geq 3082$), Proposition 10.4 follows. $\square$

For a listing of the 562 and of the 224 sporadic terms, see `http://github.com/nhf216/thesis/N2N2_41_2_Sporadic.txt`. Interestingly, these terms contain within them some other short alternations, and there are some alternations before index 2639 that "almost" follow the pattern. (Also, the $\lambda$ and $K$ both odd case never happens, as the sequence cycles through the other three cases.) For a plot of $Q_{41,2}$, see Figure 10.7.

To conclude this section, we shall consider what happens if we reverse the order of the 2 and the $N$ in the generic initial condition. This result is easier than the previous one; all sequences are eventually quasilinear.

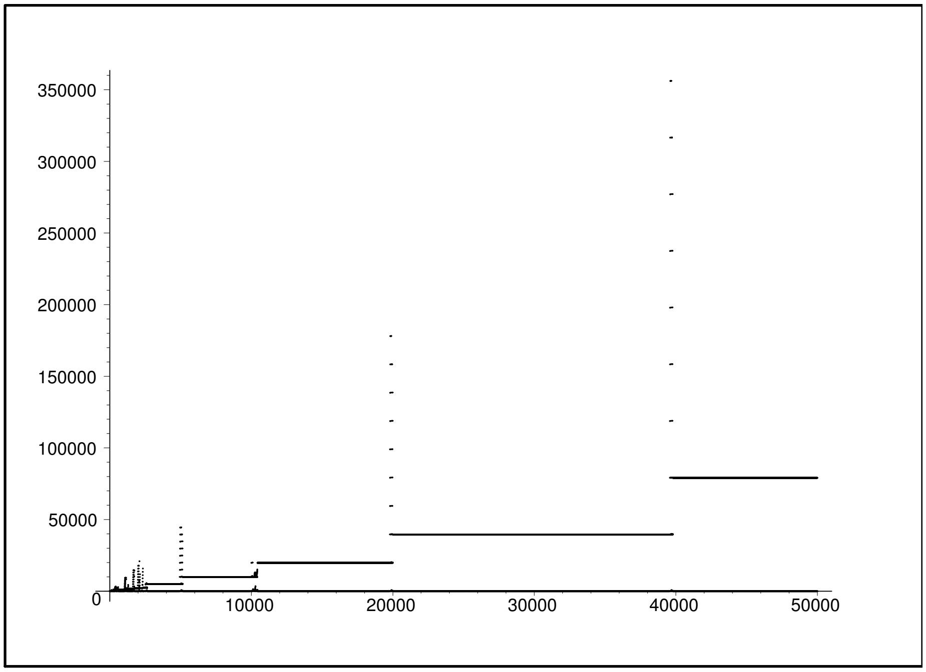


Figure 10.7: First 50000 terms of $Q_{41,2}$ (A283895)

**Theorem 10.5.** *Let $N$ be a natural number. Let $Q_{2,N}$ be the sequence resulting from the Hofstadter $Q$-recurrence and the initial condition $\langle 2, N\rangle$. We have the following cases:*

- *If $N = 1$, then we have a quasilinear solution* [31, A284429]*:*
  - $Q_{2,N}(3k) = 3$
  - $Q_{2,N}(3k + 1) = 3k + 2$
  - $Q_{2,N}(3k + 2) = 1$
- *If $N = 2$, then we have a quasilinear solution* [31, A275365]*:*
  - $Q_{2,N}(2k) = 2$
  - $Q_{2,N}(2k + 1) = 2k + 2$
- *If $N = 3$, then we have a quasilinear solution* [31, A141310]*:*
  - $Q_{2,N}(2k) = 2k + 1$
  - $Q_{2,N}(2k + 1) = 2$

- *If $N \geq 5$ is odd, then we have an eventually quasilinear solution:*
  - $Q_{2,N}(2k) = \max(N, 2k+1)$
  - $Q_{2,N}(2k+1) = 2$
- *If $N \geq 4$ is even, then we have an eventually quasilinear solution of period $N$. Each block of length $N$ begins with an anomalous value and then alternates between two different values. In the description of these solutions that follows, $0 \leq \ell < \frac{N}{2}$.*
  - $Q_{2,N}(1) = 2$
  - $Q_{2,N}(Nk) = Nk$
  - $Q_{2,N}(Nk+1) = 4$ *if* $k \geq 1$
  - $Q_{2,N}(Nk+2\ell) = N(k+1)$ *if* $\ell \geq 1$
  - $Q_{2,N}(Nk+2\ell+1) = 2$ *if* $\ell \geq 1$.

*Proof.* These are all simple inductive proofs, so they are all left as exercises. □

## 10.2 Fours and $N$'s

The initial conditions $\langle N, 4, N, 4\rangle$ and $\langle 4, N, 4, N\rangle$ yield more complicated behavior than the initial conditions with twos. First, we have the following result for $\langle N, 4, N, 4\rangle$.

**Theorem 10.6.** *Let $N$ be a natural number. Let $Q_{N,4}$ be the sequence resulting from the Hofstadter $Q$-recurrence and the initial condition $\langle N, 4, N, 4\rangle$. We have the following cases:*

1. *If $N \geq 11$ is odd, then $Q_{N,4}$ strongly dies after $N + 13$ terms.*
2. *If $N \geq 21$, $N \equiv 0 \pmod 4$, and $N$ is not 4 times a triangular number, then $Q_{N,4}$ strongly dies after $4\left\lfloor \frac{N+1+\sqrt{2N-7}}{2} \right\rfloor + 9$ terms.*
3. *If $N \geq 242$, $N \equiv 2 \pmod 8$, then $Q_{N,4}$ strongly dies after $12N + a$ terms, where $a = 50$ unless $N \equiv 10 \pmod{32}$, in which case $a = 58$.*
4. *If $N \geq 422$, $N \equiv 6 \pmod 8$, then $Q_{N,4}$ strongly dies after $14N + 34$ terms.*

*Proof.* We prove the cases in sequence.

**Case 1:** Suppose $N \geq 11$ is odd. We compute terms $Q_{N,4}(N+1)$ through $Q_{N,4}(N+13)$:

- $Q_{N,4}(N+1) = N+4$
- $Q_{N,4}(N+2) = 4$
- $Q_{N,4}(N+3) = 4$
- $Q_{N,4}(N+4) = 2N$
- $Q_{N,4}(N+5) = N+4$
- $Q_{N,4}(N+6) = 4$
- $Q_{N,4}(N+7) = N+4$
- $Q_{N,4}(N+8) = 2N+4$
- $Q_{N,4}(N+9) = N$
- $Q_{N,4}(N+10) = 4$
- $Q_{N,4}(N+11) = 2N+4$
- $Q_{N,4}(N+12) = 2N+4$
- $Q_{N,4}(N+13) = 0$

In computing these terms, we require $N$ odd and $N \geq 11$.

**Case 2:** Suppose $N \geq 21$, $N \equiv 0 \pmod 4$, and $N$ is not 4 times a triangular number. We can describe the entire sequence.

First for $N+1 \leq n \leq 2N$, we have the following pattern

- $Q_{N,4}(N+4k) = 4k+4$
- $Q_{N,4}(N+4k+1) = 2N$
- $Q_{N,4}(N+4k+2) = 4$
- $Q_{N,4}(N+4k+3) = N$

It is easy to verify that this pattern develops immediately at index $N+1$ and to inductively show that it persists through index $2N$. After index $2N$, $4k+4$ begins to exceed $N$, which causes the pattern to cease.

For simplicity of notation, let $D = \left\lfloor \frac{N+1+\sqrt{2N-7}}{2} \right\rfloor$. Next, for $2N+1 \leq n \leq 4D$ we have

- $Q_{N,4}(2N+4k) = 2k^2+6k+N+4$
- $Q_{N,4}(2N+4k+1) = 3N$
- $Q_{N,4}(2N+4k+2) = 4$

- $Q_{N,4}(2N+4k+3) = N$

It is easy to verify that this pattern develops immediately at index $2N+1$ and to check inductively that it persists for awhile. But, when does it end? The only term that changes is the $Q_{N,4}(2N+4k)$ term. This term is referenced when computing the two following terms. In the induction, we assume that $2N+4k+1-\left(2k^2+6k+N+4\right)$ lies in the initial condition; that is, that it is at least 1 and at most $N$. (We also assume the same thing about $2N+4k+2-\left(2k^2+6k+N+4\right)$, but this is less restrictive.) This value decreases with $k$, and eventually it will stop being positive. Solving $2k^2+6k+N+4 = 2N+4k$ gives us $k = \frac{-1+\sqrt{2N-7}}{2}$. So, for values of $k$ at most this value, the pattern persists. Once $k$ becomes larger than this, the pattern will cease. The last term in the pattern will be the 0 mod 4 case for the next $k$ value. So, the pattern lasts through term

$$2N+4+4\left\lfloor\frac{-1+\sqrt{2N-7}}{2}\right\rfloor = 4D,$$

as required.

Now, observe that

$$Q_{N,4}(4D) = 2\left(D-\frac{N}{2}\right)^2 + 6\left(D-\frac{N}{2}\right) + N + 4. \tag{10.1}$$

We compute the next 9 terms:

- $Q_{N,4}(4D+1) = 2N$
- $Q_{N,4}(4D+2) = 4$
- $Q_{N,4}(4D+3) = 2N$
- $Q_{N,4}(4D+4) = 2\left(D-\frac{N}{2}\right)^2 + 6\left(D-\frac{N}{2}\right) + N + 8$
- $Q_{N,4}(4D+5) = N$
- $Q_{N,4}(4D+6) = 4$
- $Q_{N,4}(4D+7) = 3N$
- $Q_{N,4}(4D+8) = 2\left(D-\frac{N}{2}\right)^2 + 6\left(D-\frac{N}{2}\right) + N + 8$
- $Q_{N,4}(4D+9) = 0.$

Most of these calculations are fairly routine, but we nevertheless illustrate a couple of them. First, we calculate the initial term.

$$\begin{aligned} Q_{N,4}(4D+1) &= Q_{N,4}(4D+1-Q_{N,4}(4D)) + Q_{N,4}(4D+1-Q_{N,4}(4D-1)) \\ &= 0 + Q_{N,4}(4D+1-N). \end{aligned}$$

We have that $N+1 \leq 4D+1-N \leq 2N$, so this equals $2N$, as required. Calculation of the next term is similar, though it depends on having $Q_{N,4}(4D) \geq 4D+2$. The only way this would not happen is if

$$Q_{N,4}(4D) = 4D+1. \tag{10.2}$$

Using Equation (10.1), we can solve (10.2) for $D$. Doing so indicates that we have equality if and only if $D$ equals

$$\frac{N-1+\sqrt{2N-5}}{2}. \tag{10.3}$$

We know that $D$ is an integer. But, since $N \equiv 0 \pmod 4$, $2N-5 \equiv 3 \pmod 4$. So, $2N-5$ cannot be a perfect square, and, as a result, (10.3) is not an integer. Hence, it is not equal to $D$. Therefore, $Q_{N,4}(4D) \geq 4D+2$, as required.

The computations through $Q_{N,4}(4D+8)$ are similar to the two we have illustrated (noting that the expression that equals $Q_{N,4}(4D+4)$ and $Q_{N,4}(4D+8)$ is equal to $Q_{N,4}(4D)+4$). We now attempt to compute $Q_{N,4}(4D+9)$. We definitely have that $3N \geq 4D+9$. If we also have

$$2\left(D-\frac{N}{2}\right)^2 + 6\left(D-\frac{N}{2}\right) + N + 8 \geq 4D+9, \tag{10.4}$$

then we obtain $Q_{N,4}(4D+9) = 0$, and we are done (as the sequence strongly dies at this point). We have equality in (10.4) when $D$ equals

$$\frac{N-1+\sqrt{2N+3}}{2}. \tag{10.5}$$

If $D$ is greater than (10.5), then we obtain the strong death we desire.

Since $N$ is not four times a triangular number, we can write $N = 2A^2+2A+4d$ for some $1 \leq d < A+1$. So, in this notation, $2A+1 < \sqrt{2N+1} < 2A+2$, so Expression (10.5) is between $A^2+A-\frac{1}{2}+A+\frac{1}{2} = A^2+2A$ and $A^2+2A+1$. Since $d \geq 1$, $D = A^2+2A+1$, and $Q_{N,4}$ strongly dies.

**Case 3 or 4:** These cases are proved by using the code in `nonstdhof.txt` to generate all of the terms. This proof takes about a day of computation time, since verifying individual cases is a fairly slow process. When trying to prove these cases, the code ends up requiring $N$ mod higher powers of 2, which means there are more cases to check than cases that appear in the theorem. Also, some of the quasipolynomials that appear within the sequence have high degree, high period, and many initial sporadic terms. These take a long time to find. As a result, the code needed to check 48 cases, corresponding to each integer between 0 and 128 that is congruent to 2 mod 8 and each integer between 0 and 256 that is congruent to 6 mod 8. The outputs of all of these cases (describing every term of each of these sequences) can be found in the directory `http://github.com/nhf216/thesis/N4N4`. A formatted version of the 2 mod 128 case can be found in Appendix E.

The computer only proves case 3 for $N \geq 315$ and case 4 for $N \geq 543$. To obtain the (tight) values of $N \geq 242$ for case 3 and $N \geq 422$ for case 4, it suffices to check the finite number of additional cases, which is done easily.

□

Figure 10.8 is a plot of the entirety of $Q_{400,4}$, which is described by case 2 of Theorem 10.6. Figures 10.9 and 10.10 are plots of $Q_{770,4}$, which is described by case 3 of Theorem 10.6.

Conspicuously absent from Theorem 10.6 is a case where $N$ is four times a triangular number. For each such $N$ we have checked individually, $Q_{N,4}$ strongly dies. But, these values of $N$ seem to lead to the longest-persisting $Q_{N,4}$ sequences. See Figures 10.11 and 10.12 for plots of $Q_{312,4}$, which persists significantly longer than any $Q_{N,4}$ described by Theorem 10.6 with $N < 312$.

Finally, we describe the sequence behaviors resulting from the initial condition $\langle 4, N, 4, N \rangle$.

**Theorem 10.7.** *Let $N$ be a natural number. Let $Q_{4,N}$ be the sequence resulting from the Hofstadter $Q$-recurrence and the initial condition $\langle 4, N, 4, N \rangle$. We have the following cases:*

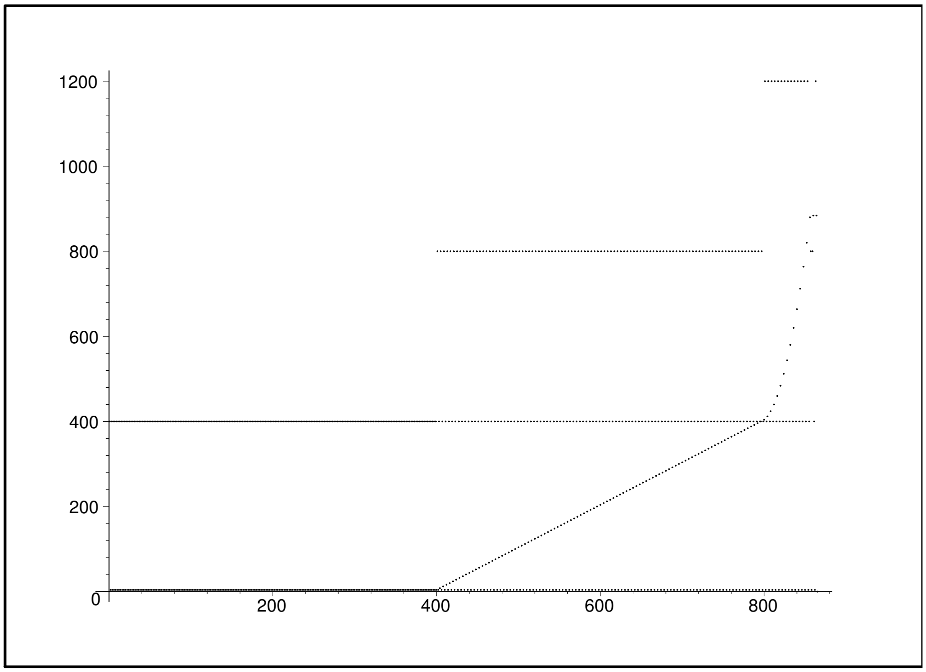


Figure 10.8: All 865 terms of sequence $Q_{400,4}$ (A283899)

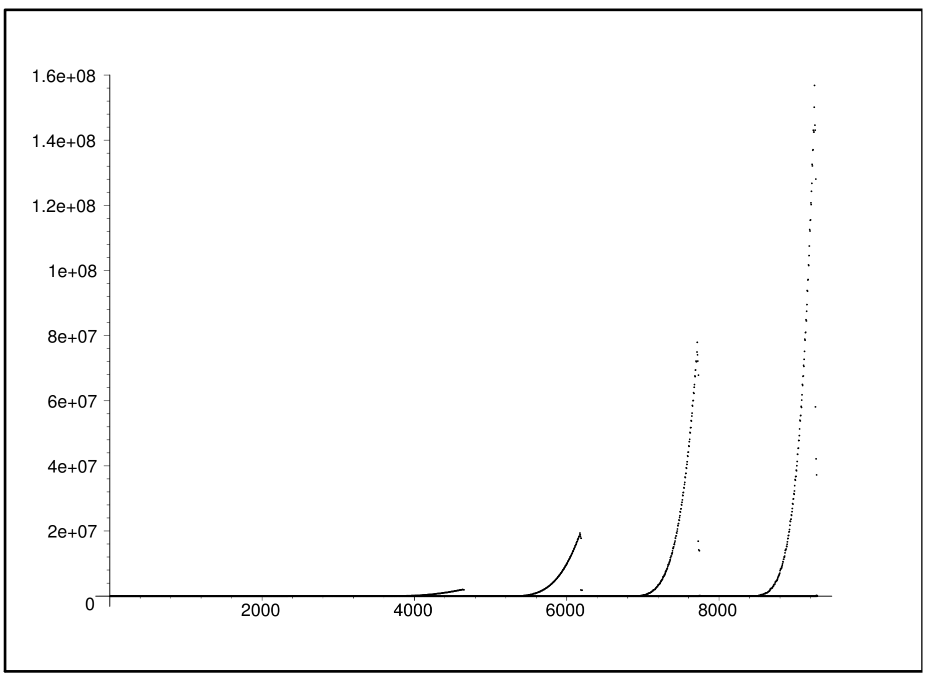


Figure 10.9: All 9290 terms of sequence $Q_{770,4}$ (A283900)

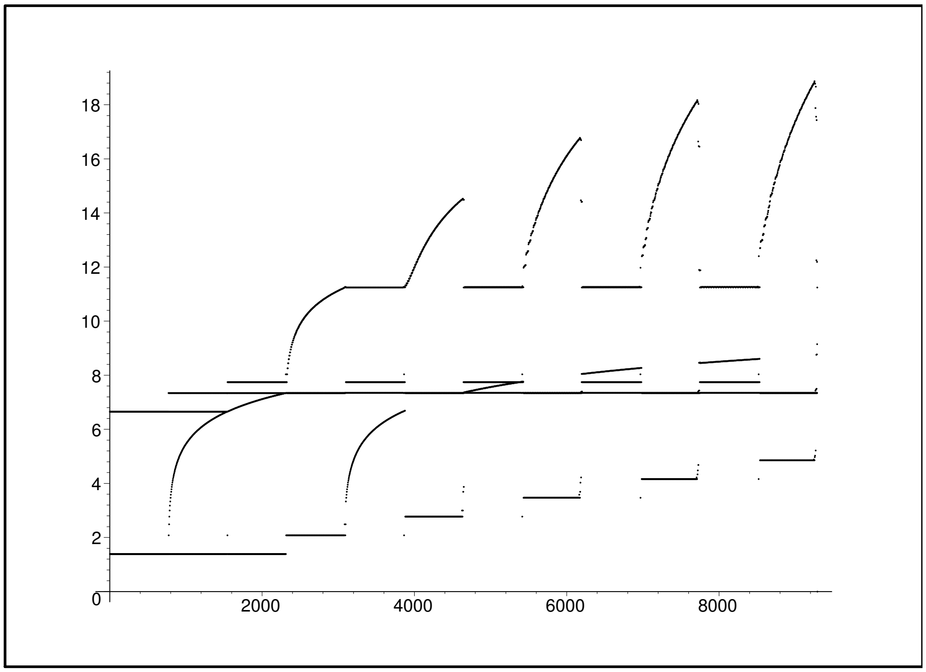


Figure 10.10: All 9290 terms of sequence $Q_{770,4}$ (A283900, log plot)

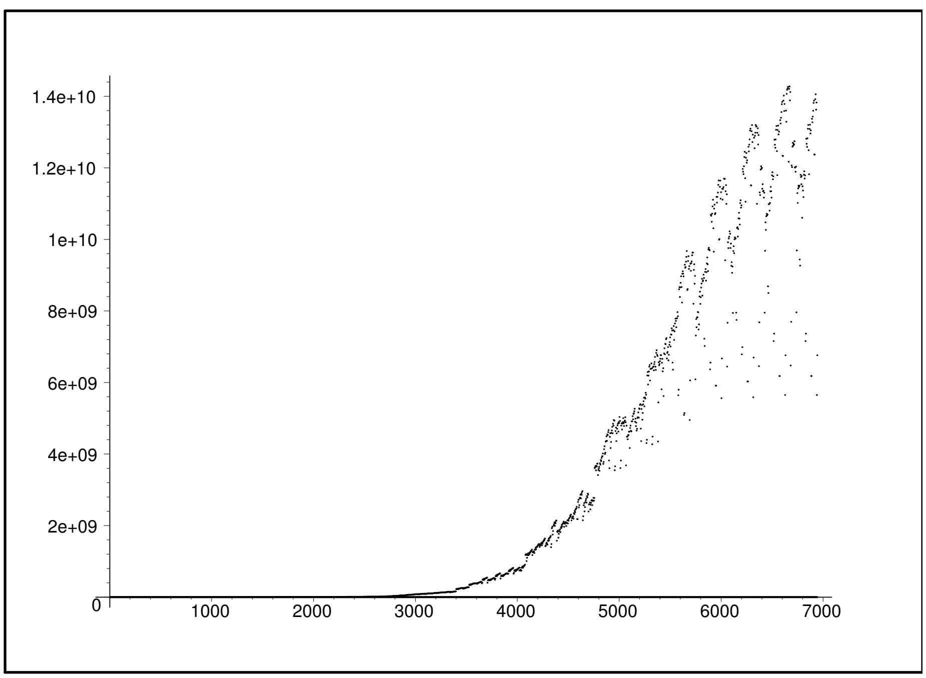


Figure 10.11: All 6944 terms of sequence $Q_{312,4}$ (A283898)

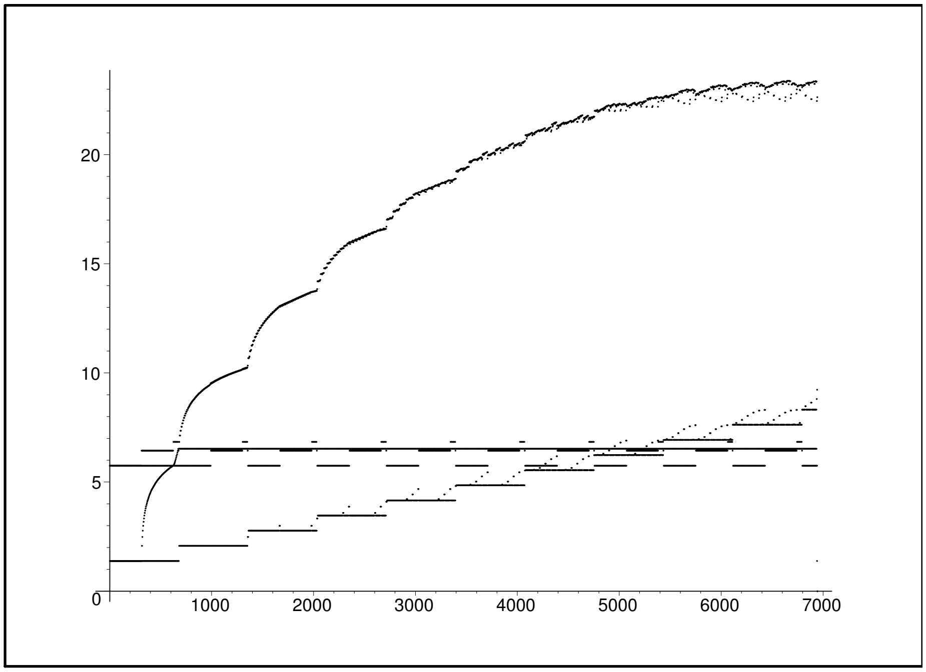


Figure 10.12: All 6944 terms of sequence $Q_{312,4}$ (A283898, log plot)

1. *If $N \geq 26$ and $N \equiv 1 \pmod 4$, then $Q_{4,N}$ dies after $2N + 28$ terms.*

2. *If $N \geq 33$ and $N \equiv 3 \pmod 4$, then $Q_{4,N}$ dies after $3N + 36$ terms.*

3. *If $N \geq 19$ and $N \equiv 0 \pmod 4$ and $N$ is not four times one more than a triangular number, then $Q_{4,N}$ dies after $4\left\lfloor \frac{N+1+\sqrt{2N-13}}{2} \right\rfloor + 6$ terms.*

*Proof.* All three of these cases are proved analogously to Theorem 10.6. The first two cases have period-4 quasilinear components of length approximately $N$ prior to their deaths. The proof of the third case is similar to the proof of case 2 in Theorem 10.6. $\square$

Figure 10.13 is a plot of all 969 terms of $Q_{4,311}$, which is described by case 2 of Theorem 10.7.

Theorem 10.7 has two notable omissions. The case of $N = 2A^2 + 2A + 4$ for some $A$ (i.e. $N$ is four times one more than a triangular number) in $Q_{4,N}$ seems to behave similarly to $Q_{N,4}$ when $N$ is four times a triangular number. The other omission is $N \equiv 2 \pmod 4$. This case seems somewhat complicated, though perhaps less complicated than the other omitted case. Attempts to automatically prove a version

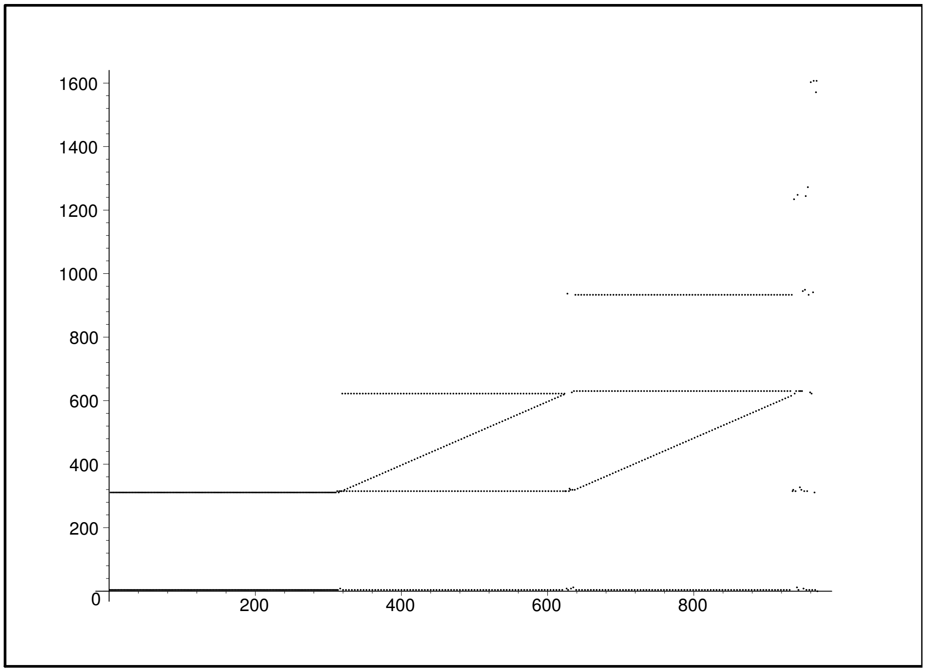


Figure 10.13: All 969 terms of sequence $Q_{4,311}$ (A283901)

of the theorem here run into issues requiring additional congruence properties of $N$, which usually means something chaotic is happening. But, plots of many of these sequences look similar. See Figures 10.14 and 10.15 for plots of $Q_{4,702}$, which provides a representative example of what these sequences typically look like.

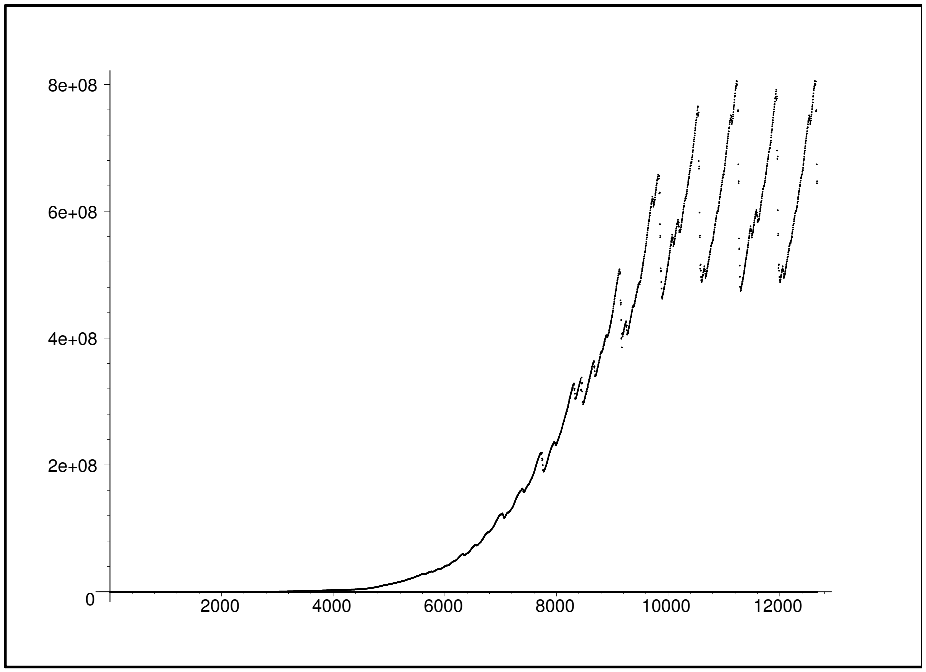


Figure 10.14: All 12671 terms of sequence $Q_{4,702}$ (A283902)

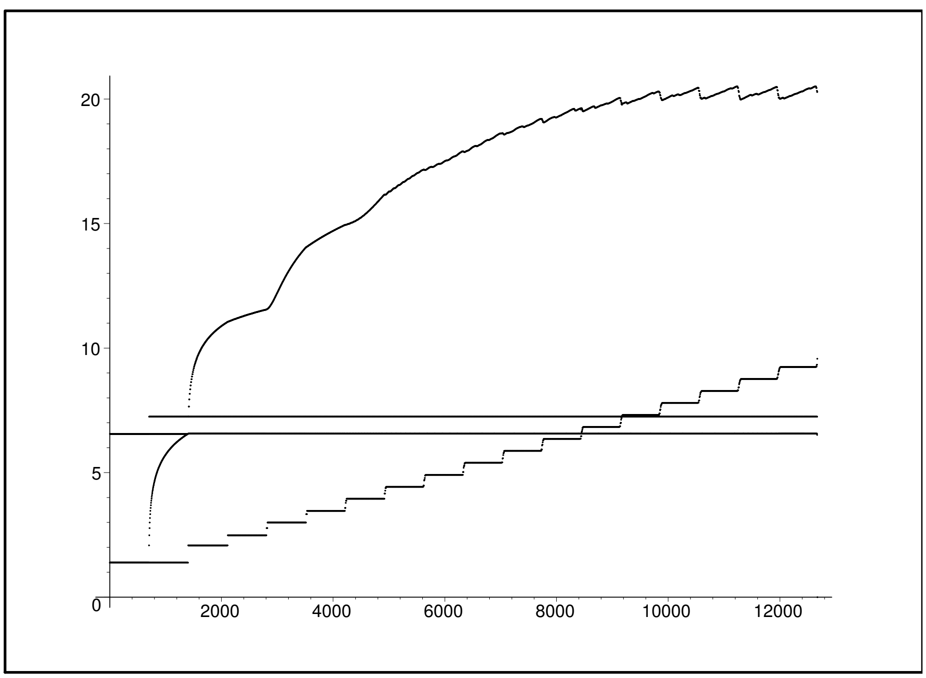


Figure 10.15: All 12671 terms of sequence $Q_{4,702}$ (A283902, log plot)

# Chapter 11

# Summary of Open Problems and Conjectures

There is an incredible diversity of open problems in the study of nested recurrences. Many such questions have been referenced in previous chapters. This chapter summarizes some of those while mentioning some additional open problems and extensions of the work of this dissertation.

The primary and most classical open question about nested recurrences is whether the Hofstadter $Q$-sequence weakly dies. The sequences analyzed in this dissertation appear to behave quite differently from the Hofstadter $Q$-sequence. The latter appears to behave largely chaotically with some hints of structure. The sequences we successfully study, on the other hand, have either an interleaved or slow structure. But, it is theoretically possible that, after a *very* long initial condition, the Hofstadter $Q$-sequence eventually becomes interleaved or slow. We consider this an unlikely, but remote, possibility.

We also mention in Chapter 2 that, throughout this dissertation, we only consider explicit solutions to recurrences. It may also make sense to study *implicit* solutions. It is unclear whether considering implicit solutions would lead to any additional results, but nobody appears to have examined them before. In particular, such solutions could potentially include infinitely many nonpositive terms.

Now, recall the main algorithm from Chapter 3. In that chapter, we specifically demand a a *basic* recurrence as input to the algorithm. In 3.2.5, we discuss the implications of extending the allowable inputs to include nonbasic linear recurrences. In practice, this generally works, though the analysis of the algorithm is no longer entirely valid. It also makes perfect sense to look for solutions to *nonlinear* nested recurrences that consist of simpler interleaved sequences. Most of the steps of our algorithm still work when the recurrences is nonlinear. The main thing that must change significantly in this setting is the algorithm for determining the orders of growth of the subsequences (Section 3.1). Nonlinearity introduces some extra complications. For example, the recurrence $D(n) = D(n - D(n-1))^2 - 2D(n - D(n-1)) + 2$ has constant solutions (constant 1 and constant 2), but it also has a solution

$$\begin{cases} D(2k) = 2^{2^{k-1}} + 1 \\ D(2k+1) = 2. \end{cases}$$

The distinction here comes from the fact that repeated squaring causes numbers to rapidly grow, unless the initial number was 0 or 1. Such concerns do not arise when dealing with only linear recurrences. Hence, it would be useful to have a version of the algorithm that can handle nonlinear recurrences. The example here includes a doubly exponential subsequence, which cannot possibly be a component of a positive-recurrent sequence. So, we would have to modify precisely what we are looking for.

Recall that solutions to the Hofstadter $Q$-recurrence are invariant under shifting (Proposition 2.7). This observation allows us to reduce our search space when looking for interleaved solutions to the $Q$-recurrence. But, even after modding out by this equivalence, we are often left with solutions that resemble each other. For example, the two period-3 solution families (modulo shifting) to the $Q$-recurrence including two constant subsequences and one linear subsequence seem like “reverses” of each other, in a sense. (See OEIS sequences A264756 and A283878.) Similarly, there are a huge number of solutions of even period that alternate between constant and standard linear subsequences. This is not particularly unexpected, as such solution families will include the sequences in the period-2 family. But these solutions currently take a lot of computational power to analyze. Perhaps we can find more symmetries between solutions to make this search faster.

Starting in Chapter 8, we begin to see nested-recurrent sequences with interleaved patterns that last awhile but not forever. We have been discovering these by generating the sequence from the initial condition and then observing the pattern. But, it should be possible to search for temporary interleaved solutions like how we look for permanent ones in Chapter 3. Perhaps the algorithm from that chapter can somehow be adapted to discover non-permanent solutions.

We now shift discussion away from the Chapter 3 algorithm. The majority of the literature on nested recurrences discusses slow solutions. It would be wonderful to have a method of automatically finding (with proof) slow solutions to nested recurrences. Unfortunately, proofs of slowness are typically considerably more involved and ad hoc than proofs of interleaved structure. The code in `http://github.com/nhf216/thesis/slowsearch.txt` can find apparent slow solutions, and there appear to be many (see `http://github.com/nhf216/thesis/slowseqs.txt`). An exploration of some of these findings may prove fruitful.

In Chapter 7, we introduce the enigmatic $R$, $S$, and $T$ sequences. We explain how to embed them into solutions to the Hofstadter $Q$-recurrence along with constant subsequences. Further, we give an initial speculation about generalizing the $R$, $S$, and $T$ sequences. Our discussion on this matter is quite preliminary; there is much more to be studied in this area.

The rest of this thesis relates to parametrized initial conditions. The first and most tantalizing of these problems is to characterize $j(N)$ and $C_j(N)$ in Theorem 9.3. The tree in Figure 9.2 has some tantalizing patterns, yet it has defied attempts at a non-recursive characterization and appears somewhat random. Perhaps further work with 5-adic numbers would be helpful in this analysis.

The end of Section 9.3 discusses the complicated eventual behaviors of the $G_{d,N}$ sequences. Perhaps it is possible to determine which $d$ values lead to long quasilinear pieces, and perhaps these sequences are less chaotic than they appear. Also, there is an infinitude of other parametrized initial condition families to study. It seems almost certain that exciting new behaviors will result from some unexplored families of initial conditions. In order to sufficiently study some of these, it may be necessary to modify the code in `http://github.com/nhf216/thesis/nonstdhof.txt` to automatically search for higher-order patterns. See the discussion after Theorem 10.1 on p. 149 for more discussion on higher-order patterns.

Many of our solutions follow a pattern for awhile, only for the pattern to break.

But, sometimes such sequences return to a similar pattern a short time later. For example, Theorem 9.3 gives us sequences consisting of arbitrarily many copies of patterns described by Lemma 9.4. Also, in Subsection 9.2.2, we observe that $B_{193}$ and $B_{3442}$ contain infinitely many Proposition 9.10 patterns, and, in Section 10.1, we observe similar behavior in the sequences $Q_{2,5}$, $Q_{2,17}$, and $Q_{2,41}$.

In general, proofs of behavior of this sort have the following ingredients:

- A parametrized description of a specific type of temporary solution. (See, for example, Proposition 9.10.)
- An argument explaining how one generic instance of said pattern transitions into another. (See, for example, the proof of Proposition 9.11.)
- A claim that the sequence in question contains a generic instance of said pattern. (See, for example, the end of the proof of Proposition 9.11.)

Theoretically, it might be possibly to automatically identify solutions of this type and to prove closed forms for them. The Maple procedures used to prove the aforementioned results are a step in that direction, but each one is specifically tailored to one specific problem. It would be worthwhile to search for more solutions like these and to develop tools to easily handle them.

Finally, and somewhat more exotically, we can tweak the strong death convention. Throughout this dissertation, we have chosen to conform with Ruskey's convention [30] and make the values at nonpositive indices of our sequences zero. But, technically, this choice is arbitrary. We could assign those indices whatever values we desire. Other conventions here could lead to additional results. The code in `http://github.com/nhf216/thesis/nicehof.txt` and `http://github.com/nhf216/thesis/nonstdhof.txt` allows the user to consider these possibilities, and a preliminary exploration indicates that this may be an untapped goldmine of additional theorems.

# Appendix A

# Some Infinite Families of Nice Solutions to the Hofstadter $Q$-Recurrence

## A.1 Period 2

There is 1 solution family to the Hofstadter $Q$-recurrence with period 2, modulo shifting.

### Solution #1

**Linear Coefficients**

$\lambda_0 = 0$ (Constant)

$\lambda_1 = 2$ (Standard Linear)

**Congruence Constraints**

$\mu_0 \equiv 0 \pmod 2$

**Equality Constraints**

$Q_q(2-\mu_1) + \mu_0 = \mu_0$

$-\mu_0 + \mu_1 + Q_q(3-\mu_1) = \mu_1$

**Inequality Constraints**

$1 \leq \mu_0$

**Formulas**

$Q_q(2k) = Q_q(2-\mu_1) + \mu_0$

$Q_q(2k+1) = 2k - \mu_0 + \mu_1 + Q_q(3-\mu_1)$

**Sample $\mu$ Values**

$\mu_0 = 2$

$\mu_1 = 0$

**Sample $Q_q$ Values**

$Q_q(2) = 0$

$Q_q(3) = 2$

**Sample IC Constraints**

$Q_q(2) = 0$

$Q_q(3) = 2$

$Q_q(4) = 2$

**Sample Initial Condition**

$\langle Q_q(1), 0, 2, 2\rangle$

## A.2 Period 3

There are 4 solution families to the Hofstadter $Q$-recurrence with period 3, modulo shifting.

# Solution #1

**Linear Coefficients**

$\lambda_0 = 0$ (Constant)

$\lambda_1 = 3$ (Standard Linear)

$\lambda_2 = 3$ (Standard Linear)

**Congruence Constraints**

$\mu_0 \equiv 0 \pmod 3$

**Equality Constraints**

$Q_q(3 - \mu_2) + Q_q(3 - \mu_1) = \mu_0$

$Q_q(2 - \mu_1) - \mu_0 + \mu_2 = \mu_2$

$-\mu_0 + \mu_1 + Q_q(4 - \mu_2) = \mu_1$

**Inequality Constraints**

$1 \leq \mu_0$

**Formulas**

$Q_q(3k) = Q_q(3 - \mu_2) + Q_q(3 - \mu_1)$

$Q_q(3k + 1) = 3k - \mu_0 + \mu_1 + Q_q(4 - \mu_2)$

$Q_q(3k + 2) = Q_q(2 - \mu_1) + 3k - \mu_0 + \mu_2$

**Sample $\mu$ Values**

$\mu_0 = 3$

$\mu_1 = 0$

$\mu_2 = -1$

**Sample $Q_q$ Values**

$Q_q(2) = 3$

$Q_q(3) = 0$

$Q_q(4) = 3$

$Q_q(5) = 3$

**Sample IC Constraints**

$Q_q(2) = 3$

$Q_q(3) = 0$

$Q_q(4) = 3$

$Q_q(5) = 3$

$Q_q(7) = 6$

$Q_q(8) = 5$

**Sample Initial Condition**

$\langle Q_q(1), 3, 0, 3, 3, Q_q(6), 6, 5 \rangle$

# Solution #2

**Linear Coefficients**

$\lambda_0 = 0$ (Constant)

$\lambda_1 = 0$ (Constant)

$\lambda_2 = \infty$ (Exponential)

**Congruence Constraints**

$\mu_0 \equiv 0 \pmod 3$

$\mu_1 \equiv 0 \pmod 3$

**Equality Constraints**

None

**Inequality Constraints**

$1 \leq \mu_0$

$1 \leq \mu_1$

**Formulas**

$Q_q(3k) = \mu_0$

$Q_q(3k+1) = \mu_1$

$Q_q(3k+2) = Q_q(3k+2-\mu_1) + Q_q(3k+2-\mu_0)$

**Sample $\mu$ Values**

$\mu_0 = 3$

$\mu_1 = 3$ **Sample IC Constraints**

$Q_q(1) = 3$

$Q_q(3) = 3$

$-Q_q(2) \leq -4$

**Sample Initial Condition**

$\langle 3, Q_q(2), 3 \rangle$

## Solution #3

**Linear Coefficients**

$\lambda_0 = 0$ (Constant)

$\lambda_1 = 0$ (Constant)

$\lambda_2 = 3$ (Standard Linear)

**Congruence Constraints**

$\mu_0 \equiv 0 \pmod 3$

$\mu_1 \equiv 2 \pmod 3$

**Equality Constraints**

$Q_q(3-\mu_2) + \mu_1 = \mu_0$

$\mu_1 + Q_q(4-\mu_2) = \mu_1$

**Inequality Constraints**

$1 \leq \mu_0$

$1 \leq \mu_1$

**Formulas**

$Q_q(3k) = Q_q(3-\mu_2) + \mu_1$

$Q_q(3k+1) = \mu_1 + Q_q(4-\mu_2)$

$Q_q(3k+2) = 3k + \mu_2$

**Sample $\mu$ Values**

$\mu_0 = 3$

$\mu_1 = 2$

$\mu_2 = 0$

**Sample $Q_q$ Values**

$Q_q(3) = 1$

$Q_q(4) = 0$

**Sample IC Constraints**

$Q_q(3) = 1$

$Q_q(4) = 0$

$Q_q(5) = 3$

$Q_q(6) = 3$

$Q_q(7) = 2$

**Sample Initial Condition**

$\langle Q_q(1), Q_q(2), 1, 0, 3, 3, 2 \rangle$

## Solution #4

**Linear Coefficients**

$\lambda_0 = 0$ (Constant)

$\lambda_1 = 0$ (Constant)

$\lambda_2 = 3$ (Standard Linear)

**Congruence Constraints**

$\mu_0 \equiv 1 \pmod 3$

$\mu_1 \equiv 0 \pmod 3$

**Equality Constraints**

$Q_q(3 - \mu_2) + \mu_0 = \mu_0$

$\mu_0 + Q_q(4 - \mu_2) = \mu_1$

**Inequality Constraints**

$1 \leq \mu_0$

$1 \leq \mu_1$

**Formulas**

$Q_q(3k) = Q_q(3 - \mu_2) + \mu_0$

$Q_q(3k + 1) = \mu_0 + Q_q(4 - \mu_2)$

$Q_q(3k + 2) = 3k + \mu_2$

**Sample $\mu$ Values**

$\mu_0 = 1$

$\mu_1 = 3$

$\mu_2 = 0$

**Sample $Q_q$ Values**

$Q_q(3) = 0$

$Q_q(4) = 2$

**Sample IC Constraints**

$Q_q(3) = 0$

$Q_q(4) = 2$

$Q_q(5) = 3$

$Q_q(6) = 1$

**Sample Initial Condition**

$\langle Q_q(1), Q_q(2), 0, 2, 3, 1\rangle$

### A.3 Period 4

There are 5 solution families to the Hofstadter $Q$-recurrence with period 4, modulo shifting.

## Solution #1

**Linear Coefficients**

$\lambda_0 = 0$ (Constant)

$\lambda_1 = 4$ (Standard Linear)

$\lambda_2 = 0$ (Constant)

$\lambda_3 = 4$ (Standard Linear)

**Congruence Constraints**

$\mu_0 \equiv 2 \pmod 4$

$\mu_2 \equiv 2 \pmod 4$

**Equality Constraints**

$Q_q(2-\mu_1)+\mu_0=\mu_2$

$Q_q(4-\mu_3)+\mu_2=\mu_0$

$-2-\mu_0+\mu_3+Q_q(5-\mu_3)=\mu_1$

$2-\mu_2+\mu_1+Q_q(3-\mu_1)=\mu_3$

**Inequality Constraints**

$1 \leq \mu_0$

$1 \leq \mu_2$

**Formulas**

$Q_q(4k)=Q_q(4-\mu_3)+\mu_2$

$Q_q(4k+1)=4k-2-\mu_0+\mu_3+Q_q(5-\mu_3)$

$Q_q(4k+2)=Q_q(2-\mu_1)+\mu_0$

$Q_q(4k+3)=4k+2-\mu_2+\mu_1+Q_q(3-\mu_1)$

**Sample $\mu$ Values**

$\mu_0=2$

$\mu_1=0$

$\mu_2=2$

$\mu_3=0$

**Sample $Q_q$ Values**

$Q_q(2)=0$

$Q_q(3)=0$

$Q_q(4)=0$

$Q_q(5)=4$

**Sample IC Constraints**

$Q_q(2)=0$

$Q_q(3)=0$

$Q_q(4)=0$

$Q_q(5)=4$

$Q_q(6)=2$

**Sample Initial Condition**

$\langle Q_q(1),0,0,0,4,2\rangle$

## Solution #2

**Linear Coefficients**

$\lambda_0=0$ (Constant)

$\lambda_1=4$ (Standard Linear)

$\lambda_2=0$ (Constant)

$\lambda_3=4$ (Standard Linear)

**Congruence Constraints**

$\mu_0 \equiv 0 \pmod 4$

$\mu_2 \equiv 2 \pmod 4$

**Equality Constraints**

$Q_q(2-\mu_1)+\mu_2=\mu_2$

$Q_q(4-\mu_3)+\mu_2=\mu_0$

$-\mu_0+\mu_1+Q_q(5-\mu_3)=\mu_1$

$2-\mu_2+\mu_1+Q_q(3-\mu_1)=\mu_3$

**Inequality Constraints**

$1 \leq \mu_0$

$1 \leq \mu_2$

**Formulas**

$Q_q(4k)=Q_q(4-\mu_3)+\mu_2$

$Q_q(4k+1)=4k-\mu_0+\mu_1+Q_q(5-\mu_3)$

$Q_q(4k+2)=Q_q(2-\mu_1)+\mu_2$

$Q_q(4k+3) = 4k+2-\mu_2+\mu_1+Q_q(3-\mu_1)$

**Sample $\mu$ Values**

$\mu_0 = 4$

$\mu_1 = 0$

$\mu_2 = 2$

$\mu_3 = 0$

**Sample $Q_q$ Values**

$Q_q(2) = 0$

$Q_q(3) = 0$

$Q_q(4) = 2$

$Q_q(5) = 4$

**Sample IC Constraints**

$Q_q(2) = 0$

$Q_q(3) = 0$

$Q_q(4) = 2$

$Q_q(5) = 4$

$Q_q(6) = 2$

**Sample Initial Condition**

$\langle Q_q(1), 0, 0, 2, 4, 2\rangle$

## Solution #3

**Linear Coefficients**

$\lambda_0 = 0$ (Constant)

$\lambda_1 = 4$ (Standard Linear)

$\lambda_2 = 0$ (Constant)

$\lambda_3 = 4$ (Standard Linear)

**Congruence Constraints**

$\mu_0 \equiv 2 \pmod 4$

$\mu_2 \equiv 0 \pmod 4$

**Equality Constraints**

$Q_q(2-\mu_1) + \mu_0 = \mu_2$

$Q_q(4-\mu_3) + \mu_0 = \mu_0$

$-\mu_2 + \mu_3 + Q_q(3-\mu_1) = \mu_3$

$-2 - \mu_0 + \mu_3 + Q_q(5-\mu_3) = \mu_1$

**Inequality Constraints**

$1 \le \mu_0$

$1 \le \mu_2$

**Formulas**

$Q_q(4k) = Q_q(4-\mu_3) + \mu_0$

$Q_q(4k+1) = 4k-2-\mu_0+\mu_3+Q_q(5-\mu_3)$

$Q_q(4k+2) = Q_q(2-\mu_1) + \mu_0$

$Q_q(4k+3) = 4k - \mu_2 + \mu_3 + Q_q(3-\mu_1)$

**Sample $\mu$ Values**

$\mu_0 = 2$

$\mu_1 = 0$

$\mu_2 = 4$

$\mu_3 = 0$

**Sample $Q_q$ Values**

$Q_q(2) = 2$

$Q_q(3) = 4$

$Q_q(4) = 0$

$Q_q(5) = 4$

$Q_q(8) = 2$

**Sample IC Constraints**

$Q_q(2) = 2$

$Q_q(3) = 4$

$Q_q(4) = 0$

$Q_q(5) = 4$

$Q_q(7) = 4$

**Sample Initial Condition**

$\langle Q_q(1), 2, 4, 0, 4, Q_q(6), 4, 2\rangle$

## Solution #4

**Linear Coefficients**

$\lambda_0 = 0$ (Constant)

$\lambda_1 = 0$ (Constant)

$\lambda_2 = 4$ (Standard Linear)

$\lambda_3 = 4$ (Standard Linear)

**Congruence Constraints**

$\mu_0 \equiv 1 \pmod 4$

$\mu_1 \equiv 0 \pmod 4$

**Equality Constraints**

$Q_q(4 - \mu_3) + Q_q(4 - \mu_2) = \mu_0$

$\mu_0 + Q_q(5 - \mu_3) = \mu_1$

$Q_q(3 - \mu_2) - \mu_1 + \mu_3 = \mu_3$

**Inequality Constraints**

$1 \leq \mu_0$

$1 \leq \mu_1$

**Formulas**

$Q_q(4k) = Q_q(4 - \mu_3) + Q_q(4 - \mu_2)$

$Q_q(4k + 1) = \mu_0 + Q_q(5 - \mu_3)$

$Q_q(4k + 2) = 4k + \mu_2$

$Q_q(4k + 3) = Q_q(3 - \mu_2) + 4k - \mu_1 + \mu_3$

**Sample $\mu$ Values**

$\mu_0 = 1$

$\mu_1 = 4$

$\mu_2 = 0$

$\mu_3 = -1$

**Sample $Q_q$ Values**

$Q_q(3) = 4$

$Q_q(4) = 1$

$Q_q(5) = 0$

$Q_q(6) = 3$

**Sample IC Constraints**

$Q_q(3) = 4$

$Q_q(4) = 1$

$Q_q(5) = 0$

$Q_q(6) = 3$

$Q_q(7) = 3$

$Q_q(9) = 4$

$Q_q(10) = 8$

**Sample Initial Condition**

$\langle Q_q(1), Q_q(2), 4, 1, 0, 3, 3, Q_q(8), 4, 8\rangle$

## Solution #5

**Linear Coefficients**

$\lambda_0 = 0$ (Constant)

$\lambda_1 = 4$ (Standard Linear)

$\lambda_2 = 0$ (Constant)

$\lambda_3 = 4$ (Standard Linear)

**Congruence Constraints**

$\mu_0 \equiv 0 \pmod 4$

$\mu_2 \equiv 0 \pmod 4$

**Equality Constraints**

$Q_q(2 - \mu_1) + \mu_2 = \mu_2$

$Q_q(4 - \mu_3) + \mu_0 = \mu_0$

$-\mu_0 + \mu_1 + Q_q(5 - \mu_3) = \mu_1$

$-\mu_2 + \mu_3 + Q_q(3 - \mu_1) = \mu_3$

**Inequality Constraints**

$1 \leq \mu_0$

$1 \leq \mu_2$

**Formulas**

$Q_q(4k) = Q_q(4 - \mu_3) + \mu_0$

$Q_q(4k+1) = 4k - \mu_0 + \mu_1 + Q_q(5 - \mu_3)$

$Q_q(4k+2) = Q_q(2 - \mu_1) + \mu_2$

$Q_q(4k+3) = 4k - \mu_2 + \mu_3 + Q_q(3 - \mu_1)$

**Sample $\mu$ Values**

$\mu_0 = 4$

$\mu_1 = 0$

$\mu_2 = 4$

$\mu_3 = 0$

**Sample $Q_q$ Values**

$Q_q(2) = 0$

$Q_q(3) = 4$

$Q_q(4) = 0$

$Q_q(5) = 4$

**Sample IC Constraints**

$Q_q(2) = 0$

$Q_q(3) = 4$

$Q_q(4) = 0$

$Q_q(5) = 4$

$Q_q(6) = 4$

$Q_q(7) = 4$

$Q_q(8) = 4$

**Sample Initial Condition**

$\langle Q_q(1), 0, 4, 0, 4, 4, 4, 4 \rangle$

### A.4 Period 5

There are 7 solution families to the Hofstadter $Q$-recurrence with period 5, modulo shifting.

# Solution #1

**Linear Coefficients**

$\lambda_0 = 0$ (Constant)

$\lambda_1 = 5$ (Standard Linear)

$\lambda_2 = 0$ (Constant)

$\lambda_3 = 5$ (Standard Linear)

$\lambda_4 = 5$ (Standard Linear)

**Congruence Constraints**

$\mu_0 \equiv 2 \pmod 5$

$\mu_2 \equiv 0 \pmod 5$

**Equality Constraints**

$Q_q(2 - \mu_1) + \mu_0 = \mu_2$

$Q_q(5 - \mu_4) + Q_q(5 - \mu_3) = \mu_0$

$Q_q(4 - \mu_3) - \mu_2 + \mu_4 = \mu_4$

$-\mu_2 + \mu_3 + Q_q(3 - \mu_1) = \mu_3$

$-3 - \mu_0 + \mu_4 + Q_q(6 - \mu_4) = \mu_1$

**Inequality Constraints**

$1 \leq \mu_0$

$1 \leq \mu_2$

**Formulas**

$Q_q(5k) = Q_q(5 - \mu_4) + Q_q(5 - \mu_3)$

$Q_q(5k+1) = 5k - 3 - \mu_0 + \mu_4 + Q_q(6 - \mu_4)$

$Q_q(5k + 2) = Q_q(2 - \mu_1) + \mu_0$

$Q_q(5k + 3) = 5k - \mu_2 + \mu_3 + Q_q(3 - \mu_1)$

$Q_q(5k + 4) = Q_q(4 - \mu_3) + 5k - \mu_2 + \mu_4$

**Sample $\mu$ Values**

$\mu_0 = 2$

$\mu_1 = 0$

$\mu_2 = 5$

$\mu_3 = 1$

$\mu_4 = 0$

**Sample $Q_q$ Values**

$Q_q(2) = 3$

$Q_q(3) = 5$

$Q_q(4) = 2$

$Q_q(5) = 0$

$Q_q(6) = 5$

**Sample IC Constraints**

$Q_q(2) = 3$

$Q_q(3) = 5$

$Q_q(4) = 2$

$Q_q(5) = 0$

$Q_q(6) = 5$

$Q_q(8) = 6$

$Q_q(9) = 5$

**Sample Initial Condition**

$\langle Q_q(1), 3, 5, 2, 0, 5, Q_q(7), 6, 5\rangle$

# Solution #2

**Linear Coefficients**

$\lambda_0 = 0$ (Constant)

$\lambda_1 = 5$ (Standard Linear)

$\lambda_2 = 0$ (Constant)

$\lambda_3 = 5$ (Standard Linear)

$\lambda_4 = 5$ (Standard Linear)

**Congruence Constraints**

$\mu_0 \equiv 0 \pmod 5$

$\mu_2 \equiv 0 \pmod 5$

**Equality Constraints**

$Q_q(2 - \mu_1) + \mu_2 = \mu_2$

$Q_q(5 - \mu_4) + Q_q(5 - \mu_3) = \mu_0$

$Q_q(4 - \mu_3) - \mu_2 + \mu_4 = \mu_4$

$-\mu_0 + \mu_1 + Q_q(6 - \mu_4) = \mu_1$

$-\mu_2 + \mu_3 + Q_q(3 - \mu_1) = \mu_3$

**Inequality Constraints**

$1 \leq \mu_0$

$1 \leq \mu_2$

**Formulas**

$Q_q(5k) = Q_q(5 - \mu_4) + Q_q(5 - \mu_3)$

$Q_q(5k + 1) = 5k - \mu_0 + \mu_1 + Q_q(6 - \mu_4)$

$Q_q(5k + 2) = Q_q(2 - \mu_1) + \mu_2$

$Q_q(5k + 3) = 5k - \mu_2 + \mu_3 + Q_q(3 - \mu_1)$

$Q_q(5k + 4) = Q_q(4 - \mu_3) + 5k - \mu_2 + \mu_4$

**Sample $\mu$ Values**

$\mu_0 = 5$

$\mu_1 = 0$

$\mu_2 = 5$

$\mu_3 = 0$

$\mu_4 = -1$

**Sample $Q_q$ Values**

$Q_q(2) = 0$

$Q_q(3) = 5$

$Q_q(4) = 5$

$Q_q(5) = 5$

$Q_q(6) = 0$

$Q_q(7) = 5$

**Sample IC Constraints**

$Q_q(2) = 0$

$Q_q(3) = 5$

$Q_q(4) = 5$

$Q_q(5) = 5$

$Q_q(6) = 0$

$Q_q(7) = 5$

$Q_q(8) = 5$

$Q_q(9) = 4$

$Q_q(10) = 5$

$Q_q(11) = 10$

**Sample Initial Condition**

$\langle Q_q(1), 0, 5, 5, 5, 0, 5, 5, 4, 5, 10\rangle$

# Solution #3

**Linear Coefficients**

$\lambda_0 = 0$ (Constant)

$\lambda_1 = 0$ (Constant)

$\lambda_2 = 5$ (Standard Linear)

$\lambda_3 = 0$ (Constant)

$\lambda_4 = \infty$ (Steep Linear)

**Congruence Constraints**

$\mu_0 \equiv 0 \pmod 5$

$\mu_1 \equiv 2 \pmod 5$

$\mu_3 \equiv 0 \pmod 5$

**Equality Constraints**

$Q_q(3 - \mu_2) + \mu_1 = \mu_3$

**Inequality Constraints**

$1 \le \mu_0$

$1 \le \mu_1$

$1 \le \mu_3$

$\mu_3 \le Q_q(4 - \mu_2) - 1$

**Formulas**

$Q_q(5k) = \mu_0$

$Q_q(5k + 1) = \mu_1$

$Q_q(5k + 2) = 5k + \mu_2$

$Q_q(5k + 3) = Q_q(3 - \mu_2) + \mu_1$

$Q_q(5k + 4) = Q_q(5k + 4 - \mu_3) + Q_q(4 - \mu_2)$

**Sample $\mu$ Values**

$\mu_0 = 5$

$\mu_1 = 2$

$\mu_2 = 0$

$\mu_3 = 5$

**Sample $Q_q$ Values**

$Q_q(3) = 3$

$Q_q(4) = 6$

**Sample IC Constraints**

$Q_q(1) = 2$

$Q_q(2) = 0$

$Q_q(3) = 3$

$Q_q(4) = 6$

$Q_q(5) = 5$

**Sample Initial Condition**

$\langle 2, 0, 3, 6, 5 \rangle$

## Solution #4

**Linear Coefficients**

$\lambda_0 = 0$ (Constant)

$\lambda_1 = 0$ (Constant)

$\lambda_2 = 5$ (Standard Linear)

$\lambda_3 = 0$ (Constant)

$\lambda_4 = 5$ (Standard Linear)

**Congruence Constraints**

$\mu_0 \equiv 0 \pmod 5$

$\mu_1 \equiv 2 \pmod 5$

$\mu_3 \equiv 2 \pmod{5}$

**Equality Constraints**

$Q_q(3 - \mu_2) + \mu_1 = \mu_3$

$Q_q(5 - \mu_4) + \mu_3 = \mu_0$

$\mu_1 + Q_q(6 - \mu_4) = \mu_1$

$2 - \mu_3 + \mu_2 + Q_q(4 - \mu_2) = \mu_4$

**Inequality Constraints**

$1 \leq \mu_0$

$1 \leq \mu_1$

$1 \leq \mu_3$

**Formulas**

$Q_q(5k) = Q_q(5 - \mu_4) + \mu_3$

$Q_q(5k + 1) = \mu_1 + Q_q(6 - \mu_4)$

$Q_q(5k + 2) = 5k + \mu_2$

$Q_q(5k + 3) = Q_q(3 - \mu_2) + \mu_1$

$Q_q(5k+4) = 5k+2-\mu_3+\mu_2+Q_q(4-\mu_2)$

**Sample $\mu$ Values**

$\mu_0 = 5$

$\mu_1 = 2$

$\mu_2 = 0$

$\mu_3 = 2$

$\mu_4 = 0$

**Sample $Q_q$ Values**

$Q_q(3) = 0$

$Q_q(4) = 0$

$Q_q(5) = 3$

$Q_q(6) = 0$

**Sample IC Constraints**

$Q_q(3) = 0$

$Q_q(4) = 0$

$Q_q(5) = 3$

$Q_q(6) = 0$

$Q_q(7) = 5$

$Q_q(10) = 5$

$Q_q(11) = 2$

**Sample Initial Condition**

$\langle Q_q(1), Q_q(2), 0, 0, 3, 0, 5, Q_q(8), Q_q(9), 5, 2\rangle$

# Solution #5

**Linear Coefficients**

$\lambda_0 = 0$ (Constant)

$\lambda_1 = 0$ (Constant)

$\lambda_2 = 5$ (Standard Linear)

$\lambda_3 = 0$ (Constant)

$\lambda_4 = 5$ (Standard Linear)

**Congruence Constraints**

$\mu_0 \equiv 1 \pmod{5}$

$\mu_1 \equiv 0 \pmod{5}$

$\mu_3 \equiv 2 \pmod{5}$

**Equality Constraints**

$Q_q(3 - \mu_2) + \mu_3 = \mu_3$

$Q_q(5-\mu_4)+\mu_3=\mu_0$

$\mu_0+Q_q(6-\mu_4)=\mu_1$

$2-\mu_3+\mu_2+Q_q(4-\mu_2)=\mu_4$

**Inequality Constraints**

$1 \leq \mu_0$

$1 \leq \mu_1$

$1 \leq \mu_3$

**Formulas**

$Q_q(5k)=Q_q(5-\mu_4)+\mu_3$

$Q_q(5k+1)=\mu_0+Q_q(6-\mu_4)$

$Q_q(5k+2)=5k+\mu_2$

$Q_q(5k+3)=Q_q(3-\mu_2)+\mu_3$

$Q_q(5k+4)=5k+2-\mu_3+\mu_2+Q_q(4-\mu_2)$

**Sample $\mu$ Values**

$\mu_0=1$

$\mu_1=5$

$\mu_2=0$

$\mu_3=2$

$\mu_4=0$

**Sample $Q_q$ Values**

$Q_q(3)=0$

$Q_q(4)=0$

$Q_q(5)=-1$

$Q_q(6)=4$

**Sample IC Constraints**

$Q_q(3)=0$

$Q_q(4)=0$

$Q_q(5)=-1$

$Q_q(6)=4$

$Q_q(7)=5$

$Q_q(8)=2$

**Sample Initial Condition**

$\langle Q_q(1), Q_q(2), 0, 0, -1, 4, 5, 2\rangle$

## Solution #6

**Linear Coefficients**

$\lambda_0=0$ (Constant)

$\lambda_1=0$ (Constant)

$\lambda_2=5$ (Standard Linear)

$\lambda_3=0$ (Constant)

$\lambda_4=5$ (Standard Linear)

**Congruence Constraints**

$\mu_0 \equiv 1 \pmod 5$

$\mu_1 \equiv 0 \pmod 5$

$\mu_3 \equiv 0 \pmod 5$

**Equality Constraints**

$Q_q(3-\mu_2)+\mu_3=\mu_3$

$Q_q(5-\mu_4)+\mu_0=\mu_0$

$\mu_0+Q_q(6-\mu_4)=\mu_1$

$-\mu_3 + \mu_4 + Q_q(4 - \mu_2) = \mu_4$

**Inequality Constraints**

$1 \leq \mu_0$

$1 \leq \mu_1$

$1 \leq \mu_3$

**Formulas**

$Q_q(5k) = Q_q(5 - \mu_4) + \mu_0$

$Q_q(5k + 1) = \mu_0 + Q_q(6 - \mu_4)$

$Q_q(5k + 2) = 5k + \mu_2$

$Q_q(5k + 3) = Q_q(3 - \mu_2) + \mu_3$

$Q_q(5k + 4) = 5k - \mu_3 + \mu_4 + Q_q(4 - \mu_2)$

**Sample $\mu$ Values**

$\mu_0 = 1$

$\mu_1 = 5$

$\mu_2 = 0$

$\mu_3 = 5$

$\mu_4 = 0$

**Sample $Q_q$ Values**

$Q_q(3) = 0$

$Q_q(4) = 5$

$Q_q(5) = 0$

$Q_q(6) = 4$

**Sample IC Constraints**

$Q_q(3) = 0$

$Q_q(4) = 5$

$Q_q(5) = 0$

$Q_q(6) = 4$

$Q_q(7) = 5$

$Q_q(8) = 5$

$Q_q(9) = 5$

$Q_q(10) = 1$

**Sample Initial Condition**

$\langle Q_q(1), Q_q(2), 0, 5, 0, 4, 5, 5, 5, 1\rangle$

# Solution #7

**Linear Coefficients**

$\lambda_0 = 0$ (Constant)

$\lambda_1 = 0$ (Constant)

$\lambda_2 = 5$ (Standard Linear)

$\lambda_3 = 0$ (Constant)

$\lambda_4 = 5$ (Standard Linear)

**Congruence Constraints**

$\mu_0 \equiv 0 \pmod 5$

$\mu_1 \equiv 2 \pmod 5$

$\mu_3 \equiv 0 \pmod 5$

**Equality Constraints**

$Q_q(3 - \mu_2) + \mu_1 = \mu_3$

$Q_q(5 - \mu_4) + \mu_0 = \mu_0$

$\mu_1 + Q_q(6 - \mu_4) = \mu_1$

$-\mu_3 + \mu_4 + Q_q(4 - \mu_2) = \mu_4$

**Inequality Constraints**

$1 \leq \mu_0$

$1 \leq \mu_1$

$1 \leq \mu_3$

**Formulas**

$Q_q(5k) = Q_q(5 - \mu_4) + \mu_0$

$Q_q(5k + 1) = \mu_1 + Q_q(6 - \mu_4)$

$Q_q(5k + 2) = 5k + \mu_2$

$Q_q(5k + 3) = Q_q(3 - \mu_2) + \mu_1$

$Q_q(5k + 4) = 5k - \mu_3 + \mu_4 + Q_q(4 - \mu_2)$

**Sample $\mu$ Values**

$\mu_0 = 5$

$\mu_1 = 2$

$\mu_2 = 0$

$\mu_3 = 5$

$\mu_4 = 0$

**Sample $Q_q$ Values**

$Q_q(3) = 3$

$Q_q(4) = 5$

$Q_q(5) = 0$

$Q_q(6) = 0$

**Sample IC Constraints**

$Q_q(3) = 3$

$Q_q(4) = 5$

$Q_q(5) = 0$

$Q_q(6) = 0$

$Q_q(7) = 5$

$Q_q(9) = 5$

$Q_q(10) = 5$

$Q_q(11) = 2$

**Sample Initial Condition**

$\langle Q_q(1), Q_q(2), 3, 5, 0, 0, 5, Q_q(8), 5, 5, 2\rangle$

# Appendix B

## Calculation of 28 Symbolic Terms of $Q_N$

Assuming $N \geq 14$, these are the first 28 terms of $Q_N$ following the initial condition.

$$\begin{aligned}\mathbf{Q_N(N+1)} &= Q_N(N+1-Q_N(N)) + Q_N(N+1-Q_N(N-1)) \\ &= Q_N(N+1-(N)) + Q_N(N+1-(N-1)) \\ &= Q_N(1) + Q_N(2) = 1+2 = \mathbf{3}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+2)} &= Q_N(N+2-Q_N(N+1)) + Q_N(N+2-Q_N(N)) \\ &= Q_N(N+2-3) + Q_N(N+2-N) \\ &= Q_N(N-1) + Q_N(2) = N-1+2 = \mathbf{N+1}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+3)} &= Q_N(N+3-Q_N(N+2)) + Q_N(N+3-Q_N(N+1)) \\ &= Q_N(N+3-(N+1)) + Q_N(N+3-3) \\ &= Q_N(2) + Q_N(N) = 2+N = \mathbf{N+2}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+4)} &= Q_N(N+4-Q_N(N+3)) + Q_N(N+4-Q_N(N+2)) \\ &= Q_N(N+4-(N+2)) + Q_N(N+4-(N+1)) \\ &= Q_N(2) + Q_N(3) = 2+3 = \mathbf{5}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+5)} &= Q_N(N+5-Q_N(N+4)) + Q_N(N+5-Q_N(N+3)) \\ &= Q_N(N+5-5) + Q_N(N+5-(N+2)) \\ &= Q_N(N) + Q_N(3) = N+3 = \mathbf{N+3}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+6)} &= Q_N(N+6-Q_N(N+5))+Q_N(N+6-Q_N(N+4))\\ &= Q_N(N+6-(N+3))+Q_N(N+6-5)\\ &= Q_N(3)+Q_N(N+1)=3+3=\mathbf{6}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+7)} &= Q_N(N+7-Q_N(N+6))+Q_N(N+7-Q_N(N+5))\\ &= Q_N(N+7-6)+Q_N(N+7-(N+3))\\ &= Q_N(N+1)+Q_N(4)=3+4=\mathbf{7}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+8)} &= Q_N(N+8-Q_N(N+7))+Q_N(N+8-Q_N(N+6))\\ &= Q_N(N+8-7)+Q_N(N+8-6)\\ &= Q_N(N+1)+Q_N(N+2)=3+N+1=\mathbf{N+4}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+9)} &= Q_N(N+9-Q_N(N+8))+Q_N(N+9-Q_N(N+7))\\ &= Q_N(N+9-(N+4))+Q_N(N+9-7)\\ &= Q_N(5)+Q_N(N+2)=5+N+1=\mathbf{N+6}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+10)} &= Q_N(N+10-Q_N(N+9))+Q_N(N+10-Q_N(N+8))\\ &= Q_N(N+10-(N+6))+Q_N(N+10-(N+4))\\ &= Q_N(4)+Q_N(6)=4+6=\mathbf{10}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+11)} &= Q_N(N+11-Q_N(N+10))+Q_N(N+11-Q_N(N+9))\\ &= Q_N(N+11-10)+Q_N(N+11-(N+6))\\ &= Q_N(N+1)+Q_N(5)=3+5=\mathbf{8}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+12)} &= Q_N(N+12-Q_N(N+11))+Q_N(N+12-Q_N(N+10))\\ &= Q_N(N+12-8)+Q_N(N+12-10)\\ &= Q_N(N+4)+Q_N(N+2)=5+N+1=\mathbf{N+6}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+13)} &= Q_N(N+13-Q_N(N+12))+Q_N(N+13-Q_N(N+11))\\ &= Q_N(N+13-(N+6))+Q_N(N+13-8)\\ &= Q_N(7)+Q_N(N+5)=7+N+3=\mathbf{N+10}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+14)} &= Q_N(N+14-Q_N(N+13))+Q_N(N+14-Q_N(N+12))\\ &= Q_N(N+14-(N+10))+Q_N(N+14-(N+6))\\ &= Q_N(4)+Q_N(8)=4+8=\mathbf{12}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+15)} &= Q_N(N+15-Q_N(N+14))+Q_N(N+15-Q_N(N+13))\\ &= Q_N(N+15-12)+Q_N(N+15-(N+10))\\ &= Q_N(N+3)+Q_N(5)=N+2+5=\mathbf{N+7}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+16)} &= Q_N(N+16-Q_N(N+15))+Q_N(N+16-Q_N(N+14))\\ &= Q_N(N+16-(N+7))+Q_N(N+16-12)\\ &= Q_N(9)+Q_N(N+4)=9+5=\mathbf{14}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+17)} &= Q_N(N+17-Q_N(N+16))+Q_N(N+17-Q_N(N+15))\\ &= Q_N(N+17-14)+Q_N(N+17-(N+7))\\ &= Q_N(N+3)+Q_N(10)=N+2+10=\mathbf{N+12}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+18)} &= Q_N(N+18-Q_N(N+17))+Q_N(N+18-Q_N(N+16))\\ &= Q_N(N+18-(N+12))+Q_N(N+18-14)\\ &= Q_N(6)+Q_N(N+4)=6+5=\mathbf{11}\end{aligned}$$

$$\begin{aligned}\mathbf{Q_N(N+19)} &= Q_N(N+19-Q_N(N+18))+Q_N(N+19-Q_N(N+17))\\ &= Q_N(N+19-11)+Q_N(N+19-(N+12))\\ &= Q_N(N+8)+Q_N(7)=N+4+7=\mathbf{N+11}\end{aligned}$$

$$\begin{aligned}
\mathbf{Q_N(N+20)} &= Q_N(N+20-Q_N(N+19)) + Q_N(N+20-Q_N(N+18)) \\
&= Q_N(N+20-(N+11)) + Q_N(N+20-11) \\
&= Q_N(9) + Q_N(N+9) = 9+N+6 = \mathbf{N+15}
\end{aligned}$$

$$\begin{aligned}
\mathbf{Q_N(N+21)} &= Q_N(N+21-Q_N(N+20)) + Q_N(N+21-Q_N(N+19)) \\
&= Q_N(N+21-(N+15)) + Q_N(N+21-(N+11)) \\
&= Q_N(6) + Q_N(10) = 6+10 = \mathbf{16}
\end{aligned}$$

$$\begin{aligned}
\mathbf{Q_N(N+22)} &= Q_N(N+22-Q_N(N+21)) + Q_N(N+22-Q_N(N+20)) \\
&= Q_N(N+22-16) + Q_N(N+22-(N+15)) \\
&= Q_N(N+6) + Q_N(7) = 6+7 = \mathbf{13}
\end{aligned}$$

$$\begin{aligned}
\mathbf{Q_N(N+23)} &= Q_N(N+23-Q_N(N+22)) + Q_N(N+23-Q_N(N+21)) \\
&= Q_N(N+23-13) + Q_N(N+23-16) \\
&= Q_N(N+10) + Q_N(N+7) = 10+7 = \mathbf{17}
\end{aligned}$$

$$\begin{aligned}
\mathbf{Q_N(N+24)} &= Q_N(N+24-Q_N(N+23)) + Q_N(N+24-Q_N(N+22)) \\
&= Q_N(N+24-17) + Q_N(N+24-13) \\
&= Q_N(N+7) + Q_N(N+11) = 7+8 = \mathbf{15}
\end{aligned}$$

$$\begin{aligned}
\mathbf{Q_N(N+25)} &= Q_N(N+25-Q_N(N+24)) + Q_N(N+25-Q_N(N+23)) \\
&= Q_N(N+25-15) + Q_N(N+25-17) \\
&= Q_N(N+10) + Q_N(N+8) = 10+N+4 = \mathbf{N+14}
\end{aligned}$$

$$\begin{aligned}
\mathbf{Q_N(N+26)} &= Q_N(N+26-Q_N(N+25)) + Q_N(N+26-Q_N(N+24)) \\
&= Q_N(N+26-(N+14)) + Q_N(N+26-15) \\
&= Q_N(12) + Q_N(N+11) = 12+8 = \mathbf{20}
\end{aligned}$$

$$\begin{aligned}
\mathbf{Q_N(N+27)} &= Q_N(N+27-Q_N(N+26)) + Q_N(N+27-Q_N(N+25)) \\
&= Q_N(N+27-20) + Q_N(N+27-(N+14)) \\
&= Q_N(N+7) + Q_N(13) = 7 + 13 = \mathbf{20}
\end{aligned}$$

$$\begin{aligned}
\mathbf{Q_N(N+28)} &= Q_N(N+28-Q_N(N+27)) + Q_N(N+28-Q_N(N+26)) \\
&= Q_N(N+28-20) + Q_N(N+28-20) \\
&= Q_N(N+8) + Q_N(N+8) = N+4+N+4 = \mathbf{2N+8}.
\end{aligned}$$

# Appendix C

# Plots of Anomalistic Living Solutions to the $Q$-recurrence

This appendix contains plots of the sequences $Q_3(= Q_2)$, $Q_4$, $Q_5$, $Q_6$, $Q_7$, $Q_9$, $Q_{10}$, and $Q_{13}$ from Subsection 9.1.1 in Chapter 9.

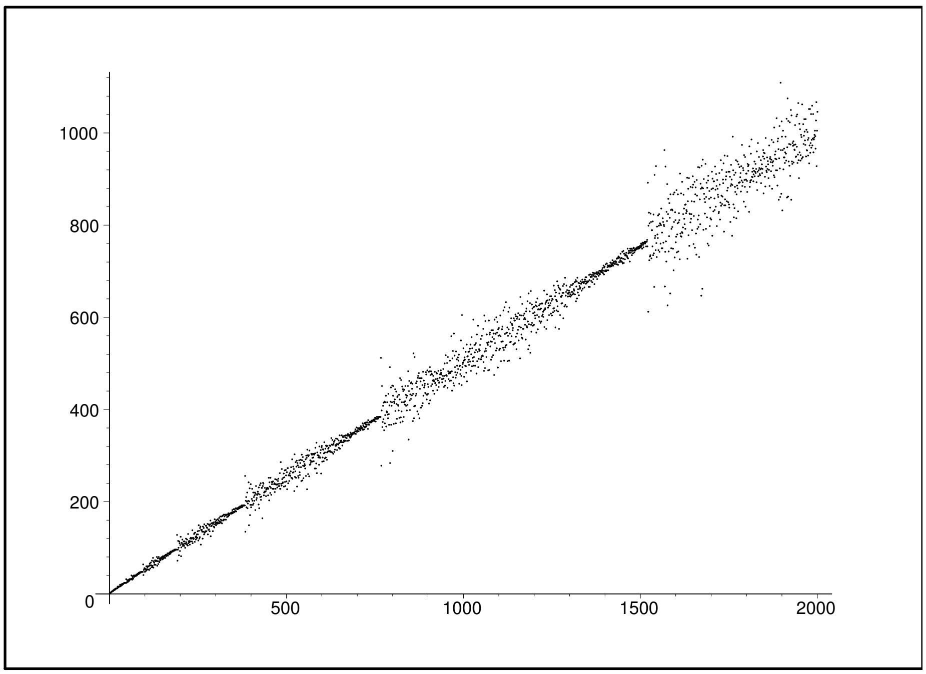


Figure C.1: The first 2000 terms of $Q_3$ (A005185 shifted)

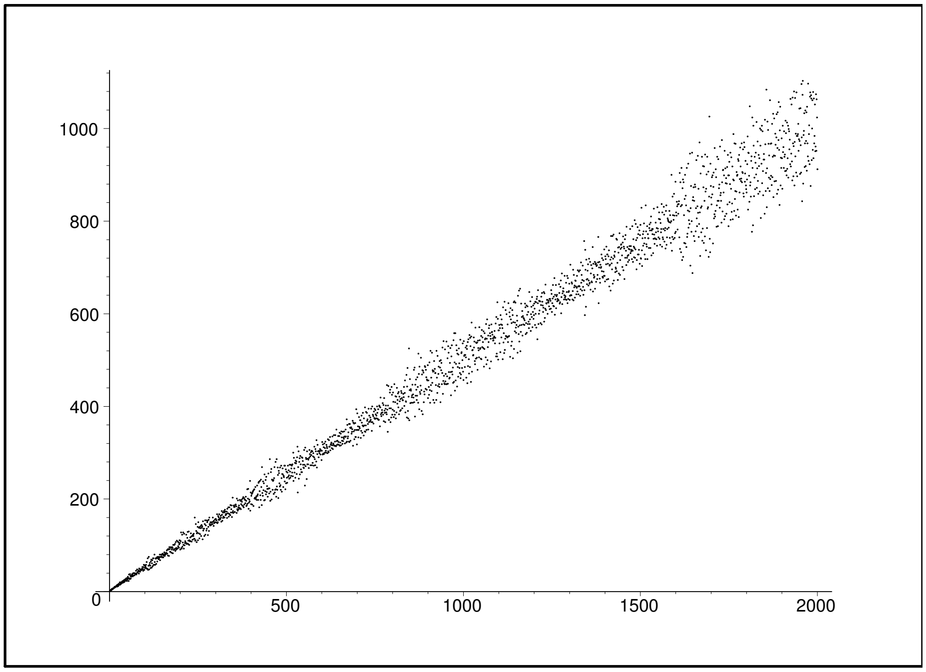


Figure C.2: The first 2000 terms of $Q_4$ (A278056)

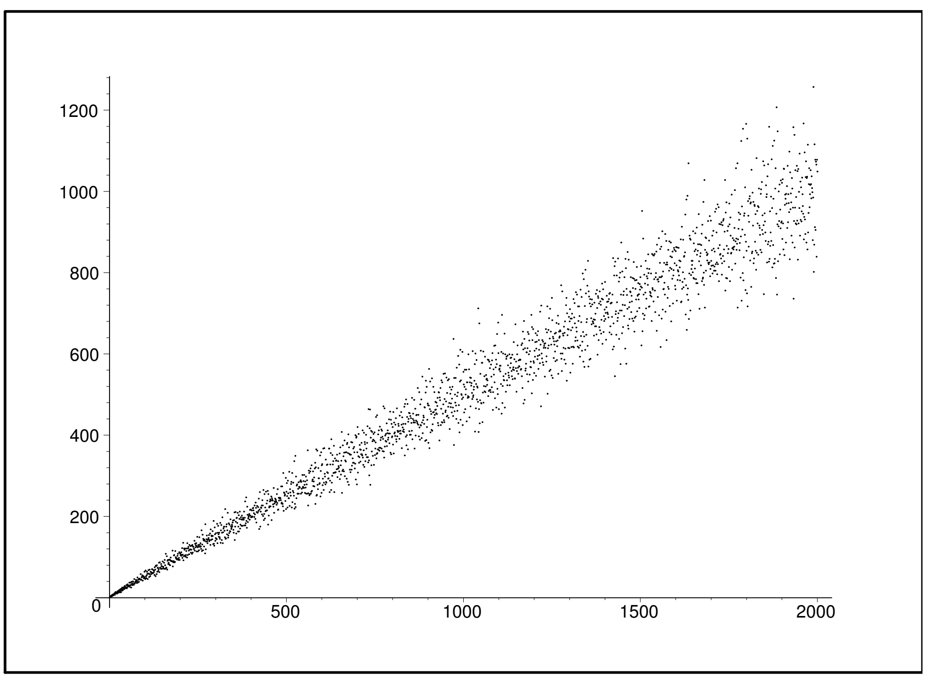


Figure C.3: The first 2000 terms of $Q_5$ (A278057)

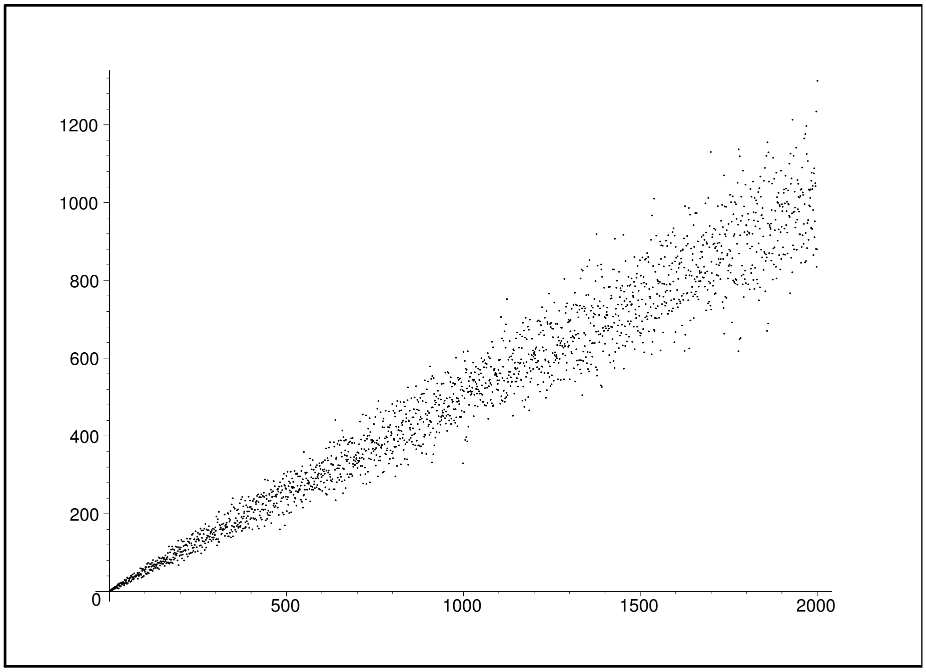


Figure C.4: The first 2000 terms of $Q_6$ (A278058)

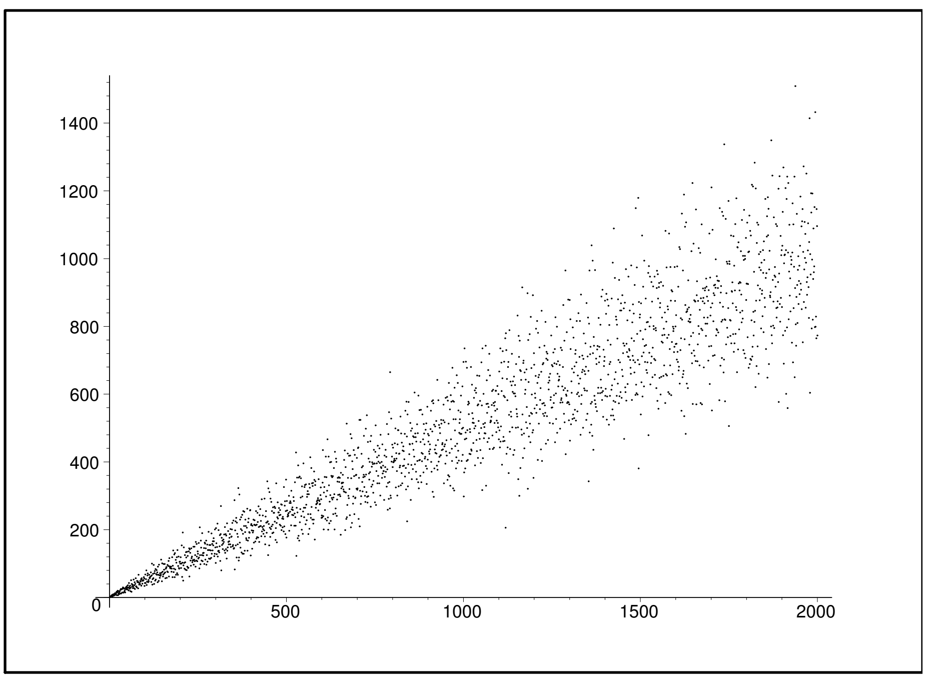


Figure C.5: The first 2000 terms of $Q_7$ (A278059)

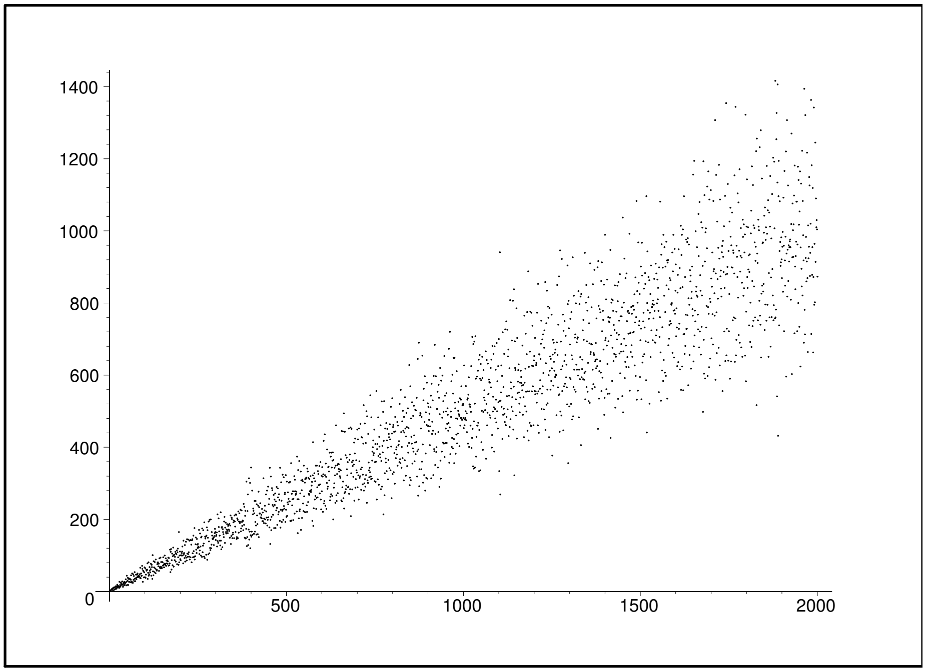


Figure C.6: The first 2000 terms of $Q_9$ (A278061)

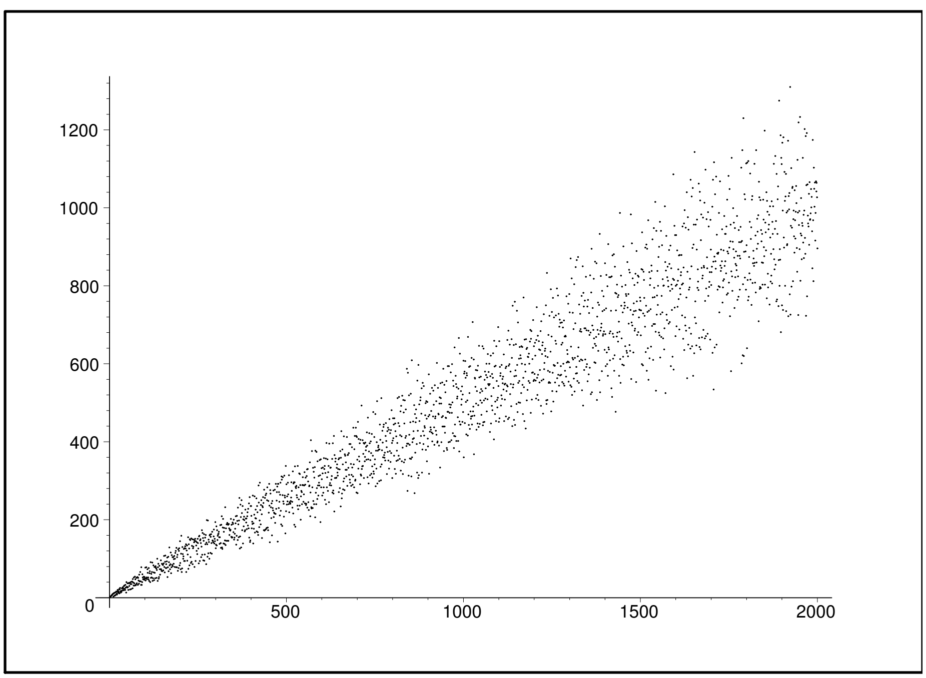


Figure C.7: The first 2000 terms of $Q_{10}$ (A278062)

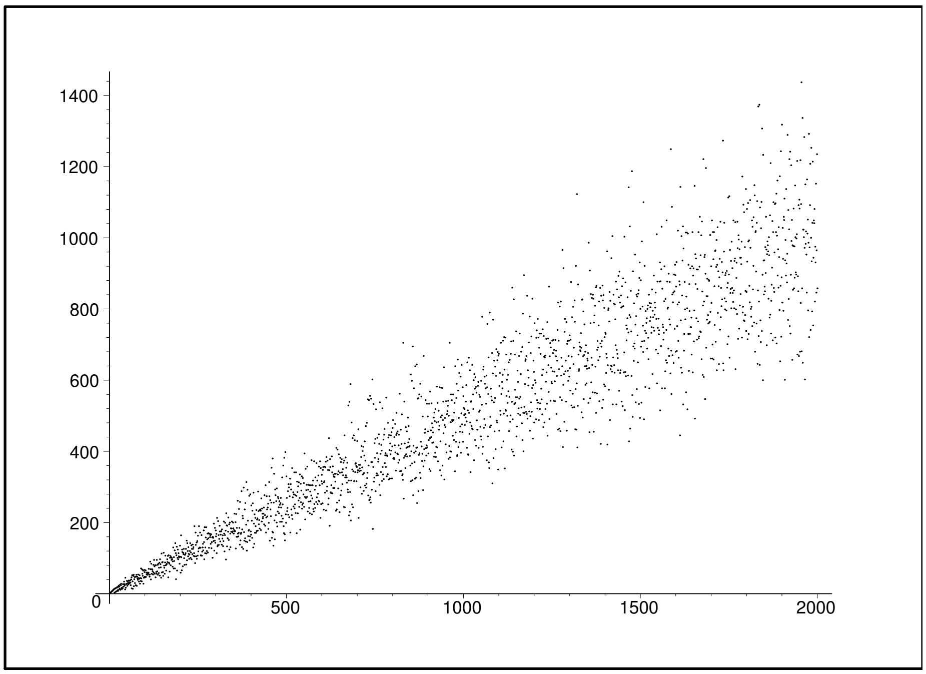


Figure C.8: The first 2000 terms of $Q_{13}$ (A278065)

# Appendix D

## The Final 158 Terms of the $C_j = 0$ Case

These are the final 158 terms in $Q_N(n)$ when $C_j = 0$ and $N \geq 118$.

- $Q_N(A_j + 3) = 6$
- $Q_N(A_j + 4) = 7$
- $Q_N(A_j + 5) = 8$
- $Q_N(A_j + 6) = 8$
- $Q_N(A_j + 7) = 10$
- $Q_N(A_j+8) = A_j \left(\frac{A_j - A_{j-1} - 2}{5}\right) + B_j + 3$
- $Q_N(A_j + 9) = 5$
- $Q_N(A_j + 10) = 8$
- $Q_N(A_j + 11) = 14$
- $Q_N(A_j + 12) = 10$
- $Q_N(A_j + 13) = 11$
- $Q_N(A_j + 14) = 13$
- $Q_N(A_j + 15) = A_j + 7$
- $Q_N(A_j + 16) = 15$
- $Q_N(A_j + 17) = A_j + 10$
- $Q_N(A_j + 18) = 14$
- $Q_N(A_j + 19) = 17$
- $Q_N(A_j + 20) = 14$
- $Q_N(A_j + 21) = 17$
- $Q_N(A_j + 22) = A_j \left(\frac{A_j - A_{j-1} - 2}{5}\right) + B_j + 11$
- $Q_N(A_j + 23) = 8$
- $Q_N(A_j + 24) = 15$
- $Q_N(A_j + 25) = A_j + 18$
- $Q_N(A_j + 26) = 22$
- $Q_N(A_j + 27) = 17$
- $Q_N(A_j + 28) = 22$
- $Q_N(A_j + 29) = 20$
- $Q_N(A_j + 30) = A_j \left(\frac{A_j - A_{j-1} - 2}{5}\right) + B_j + 11$
- $Q_N(A_j + 31) = 14$

- $Q_N(A_j+32)=14$
- $Q_N(A_j+33)=34$
- $Q_N(A_j+34) = A_j\left(\frac{A_j-A_{j-1}-2}{5}\right) + B_j+14$
- $Q_N(A_j+35)=5$
- $Q_N(A_j+36)=14$
- $Q_N(A_j+37)=22$
- $Q_N(A_j+38)=30$
- $Q_N(A_j+39)=A_j+15$
- $Q_N(A_j+40)=33$
- $Q_N(A_j+41) = A_j\left(\frac{A_j-A_{j-1}-2}{5}\right) + B_j+29$
- $Q_N(A_j+42)=5$
- $Q_N(A_j+43)=30$
- $Q_N(A_j+44)=A_j+28$
- $Q_N(A_j+45)=A_j+24$
- $Q_N(A_j+46)=40$
- $Q_N(A_j+47)=33$
- $Q_N(A_j+48) = A_j\left(\frac{A_j-A_{j-1}-2}{5}\right) + A_j+B_j+10$
- $Q_N(A_j+49)=15$
- $Q_N(A_j+50)=5$
- $Q_N(A_j+51)=54$
- $Q_N(A_j+52)=36$
- $Q_N(A_j+53)=A_j+15$
- $Q_N(A_j+54)=53$
- $Q_N(A_j+55)=A_j+40$
- $Q_N(A_j+56)=22$
- $Q_N(A_j+57)=22$
- $Q_N(A_j+58)=28$
- $Q_N(A_j+59)=36$
- $Q_N(A_j+60)=29$
- $Q_N(A_j+61)=A_j+32$
- $Q_N(A_j+62)=64$
- $Q_N(A_j+63)=36$
- $Q_N(A_j+64) = A_j\left(\frac{A_j-A_{j-1}-2}{5}\right) + B_j+22$
- $Q_N(A_j+65)=20$
- $Q_N(A_j+66)=40$
- $Q_N(A_j+67)=50$
- $Q_N(A_j+68)=36$
- $Q_N(A_j+69)=51$
- $Q_N(A_j+70) = A_j\left(\frac{A_j-A_{j-1}-2}{5}\right) + B_j+31$
- $Q_N(A_j+71)=14$
- $Q_N(A_j+72)=28$

- $Q_N(A_j + 73) = A_j + 60$
- $Q_N(A_j + 74) = 54$
- $Q_N(A_j + 75) = 32$
- $Q_N(A_j + 76) = A_j \left(\frac{A_j - A_{j-1} - 2}{5}\right) + A_j + B_j + 39$
- $Q_N(A_j + 77) = A_j + 24$
- $Q_N(A_j + 78) = 54$
- $Q_N(A_j + 79) = A_j + 73$
- $Q_N(A_j + 80) = 29$
- $Q_N(A_j + 81) = 44$
- $Q_N(A_j + 82) = A_j + 45$
- $Q_N(A_j + 83) = A_j + 53$
- $Q_N(A_j + 84) = 70$
- $Q_N(A_j + 85) = A_j + 39$
- $Q_N(A_j + 86) = 62$
- $Q_N(A_j + 87) = A_j + 66$
- $Q_N(A_j + 88) = 44$
- $Q_N(A_j + 89) = A_j + 47$
- $Q_N(A_j + 90) = 83$
- $Q_N(A_j + 91) = A_j \left(\frac{A_j - A_{j-1} - 2}{5}\right) + B_j + 47$
- $Q_N(A_j + 92) = 5$
- $Q_N(A_j + 93) = 44$
- $Q_N(A_j + 94) = A_j + 52$
- $Q_N(A_j + 95) = 97$
- $Q_N(A_j + 96) = 49$
- $Q_N(A_j + 97) = 2A_j \left(\frac{A_j - A_{j-1} - 2}{5}\right) + A_j + 2B_j + 10$
- $Q_N(A_j + 98) = 15$
- $Q_N(A_j + 99) = 70$
- $Q_N(A_j + 100) = A_j \left(\frac{A_j - A_{j-1} - 2}{5}\right) + A_j + B_j + 50$
- $Q_N(A_j + 101) = 14$
- $Q_N(A_j + 102) = 44$
- $Q_N(A_j + 103) = A_j + 83$
- $Q_N(A_j + 104) = 50$
- $Q_N(A_j + 105) = A_j + 62$
- $Q_N(A_j + 106) = 66$
- $Q_N(A_j + 107) = A_j \left(\frac{A_j - A_{j-1} - 2}{5}\right) + B_j + 74$
- $Q_N(A_j + 108) = 5$
- $Q_N(A_j + 109) = 50$
- $Q_N(A_j + 110) = A_j + 91$
- $Q_N(A_j + 111) = A_j + 52$
- $Q_N(A_j + 112) = 81$

- $Q_N(A_j + 113) = 75$
- $Q_N(A_j + 114) = A_j + 49$
- $Q_N(A_j + 115) = 99$
- $Q_N(A_j + 116) = A_j + 77$
- $Q_N(A_j + 117) = 54$
- $Q_N(A_j + 118) = A_j \left(\frac{A_j - A_{j-1} - 2}{5}\right) + B_j + 63$
- $Q_N(A_j + 119) = 20$
- $Q_N(A_j + 120) = A_j \left(\frac{A_j - A_{j-1} - 2}{5}\right) + A_j + B_j + 50$
- $Q_N(A_j + 121) = 14$
- $Q_N(A_j + 122) = 5$
- $Q_N(A_j + 123) = A_j \left(\frac{A_j - A_{j-1} - 2}{5}\right) + B_j + 113$
- $Q_N(A_j + 124) = 20$
- $Q_N(A_j + 125) = A_j + 62$
- $Q_N(A_j + 126) = 130$
- $Q_N(A_j + 127) = A_j + 65$
- $Q_N(A_j + 128) = 66$
- $Q_N(A_j + 129) = 100$
- $Q_N(A_j + 130) = 2A_j \left(\frac{A_j - A_{j-1} - 2}{5}\right) + 2B_j + 33$
- $Q_N(A_j + 131) = 14$
- $Q_N(A_j + 132) = A_j \left(\frac{A_j - A_{j-1} - 2}{5}\right) + B_j + 63$
- $Q_N(A_j + 133) = 20$
- $Q_N(A_j + 134) = A_j + 49$
- $Q_N(A_j + 135) = 185$
- $Q_N(A_j + 136) = 92$
- $Q_N(A_j + 137) = 2A_j + 24$
- $Q_N(A_j + 138) = 40$
- $Q_N(A_j + 139) = 70$
- $Q_N(A_j + 140) = 2A_j \left(\frac{A_j - A_{j-1} - 2}{5}\right) + A_j + 2B_j + 81$
- $Q_N(A_j + 141) = 14$
- $Q_N(A_j + 142) = 66$
- $Q_N(A_j + 143) = A_j + 124$
- $Q_N(A_j + 144) = 74$
- $Q_N(A_j + 145) = 35$
- $Q_N(A_j + 146) = A_j + 80$
- $Q_N(A_j + 147) = 148$
- $Q_N(A_j + 148) = A_j \left(\frac{A_j - A_{j-1} - 2}{5}\right) + B_j + 68$
- $Q_N(A_j + 149) = 5$
- $Q_N(A_j + 150) = 35$

- $Q_N(A_j + 151) = 2A_j + 157$
- $Q_N(A_j + 152) = 54$
- $Q_N(A_j + 153) = 70$
- $Q_N(A_j + 154) = A_j\left(\frac{A_j - A_{j-1} - 2}{5}\right) + A_j + B_j + 120$
- $Q_N(A_j + 155) = A_j + 39$
- $Q_N(A_j + 156) = 117$
- $Q_N(A_j + 157) = 151$
- $Q_N(A_j + 158) = A_j\left(\frac{A_j - A_{j-1} - 2}{5}\right) + B_j + 39$
- $Q_N(A_j + 159) = A_j\left(\frac{A_j - A_{j-1} - 2}{5}\right) + B_j + 3$
- $Q_N(A_j + 160) = 0$

# Appendix E

## A Complete Description of $Q_{4,N}$ when $N \equiv 2 \pmod{128}$

**For indices from $1$ to $N$:**

- $Q_{4,N}(2k) = 4$
- $Q_{4,N}(2k+1) = N$

**For indices from $N+1$ to $2N$:**

- $Q_{4,N}(N+4k) = 4k+41$
- $Q_{4,N}(N+4k+1) = 2N$
- $Q_{4,N}(N+4k+2) = 4$
- $Q_{4,N}(N+4k+3) = N$

**Some sporadic values:**

- $Q_{4,N}(2N+1) = 2N$
- $Q_{4,N}(2N+2) = N+61$
- $Q_{4,N}(2N+3) = 2N$
- $Q_{4,N}(2N+4) = 8$

**For indices from $2N+5$ to $3N$:**

- $Q_{4,N}(2N+4k) = 4$
- $Q_{4,N}(2N+4k+1) = 2N$
- $Q_{4,N}(2N+4k+2) = 4k+N+21$
- $Q_{4,N}(2N+4k+3) = 3N$

**Some sporadic values:**

- $Q_{4,N}(3N+1) = 4N$
- $Q_{4,N}(3N+2) = 4$
- $Q_{4,N}(3N+3) = 2N$
- $Q_{4,N}(3N+4) = 2N+81$
- $Q_{4,N}(3N+5) = 3N$
- $Q_{4,N}(3N+6) = 8$
- $Q_{4,N}(3N+7) = 3N$
- $Q_{4,N}(3N+8) = 2N+41$
- $Q_{4,N}(3N+9) = 3N$
- $Q_{4,N}(3N+10) = 8$
- $Q_{4,N}(3N+11) = 3N$
- $Q_{4,N}(3N+12) = 2N+121$
- $Q_{4,N}(3N+13) = 3N$

**For indices from $3N+14$ to $4N$:**

- $Q_{4,N}(3N+8k) = 2N+41$

- $Q_{4,N}(3N + 8k + 1) = 2N$
- $Q_{4,N}(3N + 8k + 2) = 8$
- $Q_{4,N}(3N + 8k + 3) = Nk + 2N$
- $Q_{4,N}(3N + 8k + 4) = 2N + 121$
- $Q_{4,N}(3N + 8k + 5) = 2N$
- $Q_{4,N}(3N + 8k + 6) = 8$
- $Q_{4,N}(3N + 8k + 7) = Nk + 3N$

**Some sporadic values:**

- $Q_{4,N}(4N + 1) = \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(4N + 2) = 2N + 121$
- $Q_{4,N}(4N + 3) = 2N$
- $Q_{4,N}(4N + 4) = 12$
- $Q_{4,N}(4N + 5) = \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(4N + 6) = 2N + 121$
- $Q_{4,N}(4N + 7) = 2N$
- $Q_{4,N}(4N + 8) = 8$
- $Q_{4,N}(4N + 9) = \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(4N + 10) = 2N + 121$
- $Q_{4,N}(4N + 11) = 2N$
- $Q_{4,N}(4N + 12) = 8$
- $Q_{4,N}(4N + 13) = \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(4N + 14) = 2N + 121$
- $Q_{4,N}(4N + 15) = 2N$
- $Q_{4,N}(4N + 16) = 12$
- $Q_{4,N}(4N + 17) = \frac{1}{8}N^2 + \frac{19}{4}N$

**For indices from $4N+18$ to $5N+5$:**

- $Q_{4,N}(4N + 4k) = 4k + 81$
- $Q_{4,N}(4N + 4k + 1) = \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(4N + 4k + 2) = 2N + 121$
- $Q_{4,N}(4N + 4k + 3) = 3N$

**Some sporadic values:**

- $Q_{4,N}(5N + 6) = N + 101$
- $Q_{4,N}(5N + 7) = \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(5N + 8) = 2N + 41$
- $Q_{4,N}(5N + 9) = 3N$
- $Q_{4,N}(5N + 10) = N + 181$
- $Q_{4,N}(5N + 11) = \frac{1}{8}N^2 + \frac{15}{4}N$
- $Q_{4,N}(5N + 12) = 2N + 121$
- $Q_{4,N}(5N + 13) = 4N$
- $Q_{4,N}(5N + 14) = 8$
- $Q_{4,N}(5N + 15) = \frac{1}{8}N^2 + \frac{23}{4}N$
- $Q_{4,N}(5N + 16) = 2N + 41$
- $Q_{4,N}(5N + 17) = 3N$
- $Q_{4,N}(5N + 18) = N + 261$
- $Q_{4,N}(5N + 19) = \frac{1}{8}N^2 + \frac{15}{4}N$

- $Q_{4,N}(5N+20) = 2N+121$
- $Q_{4,N}(5N+21) = 3N$
- $Q_{4,N}(5N+22) = N+301$
- $Q_{4,N}(5N+23) = \frac{1}{8}N^2 + \frac{15}{4}N$
- $Q_{4,N}(5N+24) = 2N+121$
- $Q_{4,N}(5N+25) = 3N$
- $Q_{4,N}(5N+26) = N+341$

**For indices from $5N+27$ to $6N$:**

- $Q_{4,N}(5N+16k) = 2N+41$
- $Q_{4,N}(5N+16k+1) = 2N$
- $Q_{4,N}(5N+16k+2) = 16$
- $Q_{4,N}(5N+16k+3) = Nk^2+3Nk+\frac{1}{8}N^2 - \frac{1}{4}N$
- $Q_{4,N}(5N+16k+4) = 2N+121$
- $Q_{4,N}(5N+16k+5) = 2N$
- $Q_{4,N}(5N+16k+6) = 16$
- $Q_{4,N}(5N+16k+7) = Nk^2+4Nk+\frac{1}{8}N^2 - \frac{5}{4}N$
- $Q_{4,N}(5N+16k+8) = 2N+121$
- $Q_{4,N}(5N+16k+9) = 2N$
- $Q_{4,N}(5N+16k+10) = 16$
- $Q_{4,N}(5N+16k+11) = Nk^2+4Nk+\frac{1}{8}N^2 - \frac{5}{4}N$
- $Q_{4,N}(5N+16k+12) = 2N+121$
- $Q_{4,N}(5N+16k+13) = 2N$
- $Q_{4,N}(5N+16k+14) = 16$
- $Q_{4,N}(5N+16k+15) = Nk^2+5Nk+\frac{1}{8}N^2 + \frac{23}{4}N$

**Some sporadic values:**

- $Q_{4,N}(6N+1) = \frac{1}{256}N^3+\frac{19}{64}N^2+\frac{25}{64}N$
- $Q_{4,N}(6N+2) = 2N+121$
- $Q_{4,N}(6N+3) = 2N$
- $Q_{4,N}(6N+4) = 20$
- $Q_{4,N}(6N+5) = \frac{1}{256}N^3+\frac{19}{64}N^2+\frac{25}{64}N$
- $Q_{4,N}(6N+6) = 2N+121$
- $Q_{4,N}(6N+7) = 2N$
- $Q_{4,N}(6N+8) = 16$
- $Q_{4,N}(6N+9) = \frac{1}{256}N^3+\frac{23}{64}N^2+\frac{17}{64}N$
- $Q_{4,N}(6N+10) = 2N+121$
- $Q_{4,N}(6N+11) = 2N$
- $Q_{4,N}(6N+12) = 16$
- $Q_{4,N}(6N+13) = \frac{1}{256}N^3 + \frac{27}{64}N^2 + \frac{393}{64}N$
- $Q_{4,N}(6N+14) = 2N+41$
- $Q_{4,N}(6N+15) = 2N$
- $Q_{4,N}(6N+16) = 20$

- $Q_{4,N}(6N+17) = \frac{1}{256}N^3 + \frac{27}{64}N^2 + \frac{393}{64}N$
- $Q_{4,N}(6N+18) = 2N+41$
- $Q_{4,N}(6N+19) = 2N$
- $Q_{4,N}(6N+20) = 40$
- $Q_{4,N}(6N+21) = \frac{1}{256}N^3 + \frac{19}{64}N^2 + \frac{153}{64}N$
- $Q_{4,N}(6N+22) = 2N+41$
- $Q_{4,N}(6N+23) = 3N$
- $Q_{4,N}(6N+24) = 2N+321$
- $Q_{4,N}(6N+25) = \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(6N+26) = 2N+121$
- $Q_{4,N}(6N+27) = 2N$
- $Q_{4,N}(6N+28) = 48$
- $Q_{4,N}(6N+29) = \frac{1}{256}N^3 + \frac{19}{64}N^2 + \frac{153}{64}N$
- $Q_{4,N}(6N+30) = 2N+41$
- $Q_{4,N}(6N+31) = 3N$
- $Q_{4,N}(6N+32) = 2N+401$
- $Q_{4,N}(6N+33) = \frac{1}{8}N^2 + \frac{11}{4}N$

**For indices from $6N+34$ to $7N+5$:**

- $Q_{4,N}(6N+8k) = 8k+2N$
- $Q_{4,N}(6N+8k+1) = \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(6N+8k+2) = 2N+121$
- $Q_{4,N}(6N+8k+3) = 3N$
- $Q_{4,N}(6N+8k+4) = 8k+2N+121$
- $Q_{4,N}(6N+8k+5) = \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(6N+8k+6) = 2N+121$
- $Q_{4,N}(6N+8k+7) = 3N$

**Some sporadic values:**

- $Q_{4,N}(7N+6) = 3N+141$
- $Q_{4,N}(7N+7) = \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(7N+8) = 2N+121$
- $Q_{4,N}(7N+9) = 3N$
- $Q_{4,N}(7N+10) = 3N+181$
- $Q_{4,N}(7N+11) = \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(7N+12) = 2N+121$
- $Q_{4,N}(7N+13) = 3N$
- $Q_{4,N}(7N+14) = 3N+221$
- $Q_{4,N}(7N+15) = \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(7N+16) = 2N+121$
- $Q_{4,N}(7N+17) = 3N$
- $Q_{4,N}(7N+18) = 3N+221$
- $Q_{4,N}(7N+19) = \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(7N+20) = 2N+41$

- $Q_{4,N}(7N+21) = 3N$
- $Q_{4,N}(7N+22) = 3N+381$
- $Q_{4,N}(7N+23) = \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(7N+24) = 2N+121$
- $Q_{4,N}(7N+25) = 4N$
- $Q_{4,N}(7N+26) = 16$
- $Q_{4,N}(7N+27) = \frac{1}{8}N^2 + \frac{31}{4}N$
- $Q_{4,N}(7N+28) = 2N+121$
- $Q_{4,N}(7N+29) = 3N$
- $Q_{4,N}(7N+30) = 3N+381$
- $Q_{4,N}(7N+31) = \frac{1}{8}N^2 + \frac{15}{4}N$
- $Q_{4,N}(7N+32) = 2N+121$
- $Q_{4,N}(7N+33) = 3N$
- $Q_{4,N}(7N+34) = 3N+421$
- $Q_{4,N}(7N+35) = \frac{1}{8}N^2 + \frac{15}{4}N$
- $Q_{4,N}(7N+36) = 2N+121$
- $Q_{4,N}(7N+37) = 3N$
- $Q_{4,N}(7N+38) = 3N+461$

**For indices from $7N+39$ to $8N$:**

- $Q_{4,N}(7N+32k) = 2N+121$
- $Q_{4,N}(7N+32k+1) = 2N$
- $Q_{4,N}(7N+32k+2) = 32$
- $Q_{4,N}(7N+32k+3) = \frac{4}{3}Nk^3+5Nk^2+\frac{1}{8}N^2k+\frac{41}{12}Nk-6N$
- $Q_{4,N}(7N+32k+4) = 2N+121$
- $Q_{4,N}(7N+32k+5) = 2N$
- $Q_{4,N}(7N+32k+6) = 32$
- $Q_{4,N}(7N+32k+7) = \frac{4}{3}Nk^3+6Nk^2+\frac{1}{8}N^2k+\frac{41}{12}Nk-7N$
- $Q_{4,N}(7N+32k+8) = 2N+121$
- $Q_{4,N}(7N+32k+9) = 2N$
- $Q_{4,N}(7N+32k+10) = 32$
- $Q_{4,N}(7N+32k+11) = \frac{4}{3}Nk^3+6Nk^2+\frac{1}{8}N^2k+\frac{41}{12}Nk+\frac{1}{8}N^2+\frac{11}{4}N$
- $Q_{4,N}(7N+32k+12) = 2N+121$
- $Q_{4,N}(7N+32k+13) = 2N$
- $Q_{4,N}(7N+32k+14) = 32$
- $Q_{4,N}(7N+32k+15) = \frac{4}{3}Nk^3+7Nk^2+\frac{1}{8}N^2k+\frac{137}{12}Nk+\frac{1}{8}N^2+\frac{11}{4}N$
- $Q_{4,N}(7N+32k+16) = 2N+121$
- $Q_{4,N}(7N+32k+17) = 2N$
- $Q_{4,N}(7N+32k+18) = 32$
- $Q_{4,N}(7N+32k+19) = \frac{4}{3}Nk^3+7Nk^2+\frac{1}{8}N^2k+\frac{113}{12}Nk+\frac{1}{8}N^2+\frac{19}{4}N$
- $Q_{4,N}(7N+32k+20) = 2N+41$
- $Q_{4,N}(7N+32k+21) = 2N$

- $Q_{4,N}(7N+32k+22) = 32$
- $Q_{4,N}(7N+32k+23) = \frac{4}{3}Nk^3 + 8Nk^2 + \frac{1}{8}N^2k + \frac{125}{12}Nk + \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(7N+32k+24) = 2N+121$
- $Q_{4,N}(7N+32k+25) = 2N$
- $Q_{4,N}(7N+32k+26) = 32$
- $Q_{4,N}(7N+32k+27) = \frac{4}{3}Nk^3 + 8Nk^2 + \frac{1}{8}N^2k + \frac{125}{12}Nk + \frac{1}{8}N^2 + \frac{31}{4}N$
- $Q_{4,N}(7N+32k+28) = 2N+121$
- $Q_{4,N}(7N+32k+29) = 2N$
- $Q_{4,N}(7N+32k+30) = 32$
- $Q_{4,N}(7N+32k+31) = \frac{4}{3}Nk^3 + 9Nk^2 + \frac{1}{8}N^2k + \frac{233}{12}Nk + \frac{1}{8}N^2 + \frac{15}{4}N$

**Some sporadic values:**

- $Q_{4,N}(8N+1) = \frac{1}{24576}N^4 + \frac{35}{4096}N^3 + \frac{491}{6144}N^2 - \frac{5319}{1024}N$
- $Q_{4,N}(8N+2) = 2N+121$
- $Q_{4,N}(8N+3) = 2N$
- $Q_{4,N}(8N+4) = 36$
- $Q_{4,N}(8N+5) = \frac{1}{24576}N^4 + \frac{35}{4096}N^3 + \frac{491}{6144}N^2 - \frac{5319}{1024}N$
- $Q_{4,N}(8N+6) = 2N+121$
- $Q_{4,N}(8N+7) = 2N$
- $Q_{4,N}(8N+8) = 32$
- $Q_{4,N}(8N+9) = \frac{1}{24576}N^4 + \frac{39}{4096}N^3 + \frac{1235}{6144}N^2 + \frac{4669}{1024}N$
- $Q_{4,N}(8N+10) = 2N+121$
- $Q_{4,N}(8N+11) = 2N$
- $Q_{4,N}(8N+12) = 32$
- $Q_{4,N}(8N+13) = \frac{1}{24576}N^4 + \frac{43}{4096}N^3 + \frac{2747}{6144}N^2 + \frac{3137}{1024}N$
- $Q_{4,N}(8N+14) = 2N+121$
- $Q_{4,N}(8N+15) = 2N$
- $Q_{4,N}(8N+16) = 40$
- $Q_{4,N}(8N+17) = \frac{1}{24576}N^4 + \frac{39}{4096}N^3 + \frac{1619}{6144}N^2 + \frac{10685}{1024}N$
- $Q_{4,N}(8N+18) = 2N+121$
- $Q_{4,N}(8N+19) = 2N$
- $Q_{4,N}(8N+20) = 56$
- $Q_{4,N}(8N+21) = \frac{1}{24576}N^4 + \frac{35}{4096}N^3 + \frac{875}{6144}N^2 - \frac{7495}{1024}N$
- $Q_{4,N}(8N+22) = 2N+121$
- $Q_{4,N}(8N+23) = 2N$
- $Q_{4,N}(8N+24) = 2N+481$
- $Q_{4,N}(8N+25) = \frac{1}{256}N^3 + \frac{15}{64}N^2 - \frac{159}{64}N$

- $Q_{4,N}(8N+26) = 2N+121$
- $Q_{4,N}(8N+27) = 2N$
- $Q_{4,N}(8N+28) = 68$
- $Q_{4,N}(8N+29) = \frac{1}{24576}N^4 + \frac{31}{4096}N^3 - \frac{61}{6144}N^2 + \frac{11253}{1024}N$
- $Q_{4,N}(8N+30) = 2N+121$
- $Q_{4,N}(8N+31) = 2N$
- $Q_{4,N}(8N+32) = 2N+801$
- $Q_{4,N}(8N+33) = \frac{1}{256}N^3 + \frac{3}{64}N^2 + \frac{185}{64}N$
- $Q_{4,N}(8N+34) = 2N+121$
- $Q_{4,N}(8N+35) = 3N$
- $Q_{4,N}(8N+36) = 4N+441$
- $Q_{4,N}(8N+37) = \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(8N+38) = 2N+121$
- $Q_{4,N}(8N+39) = 2N$
- $Q_{4,N}(8N+40) = 2N+881$
- $Q_{4,N}(8N+41) = \frac{1}{256}N^3 + \frac{3}{64}N^2 + \frac{313}{64}N$
- $Q_{4,N}(8N+42) = 2N+121$
- $Q_{4,N}(8N+43) = 3N$
- $Q_{4,N}(8N+44) = 4N+521$
- $Q_{4,N}(8N+45) = \frac{1}{8}N^2 + \frac{11}{4}N$

**For indices from $8N+46$ to $9N+21$:**

- $Q_{4,N}(8N+16k) = 16k+4N$
- $Q_{4,N}(8N+16k+1) = \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(8N+16k+2) = 2N+121$
- $Q_{4,N}(8N+16k+3) = 3N$
- $Q_{4,N}(8N+16k+4) = 16k+4N+41$
- $Q_{4,N}(8N+16k+5) = \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(8N+16k+6) = 2N+121$
- $Q_{4,N}(8N+16k+7) = 3N$
- $Q_{4,N}(8N+16k+8) = 16k+4N+161$
- $Q_{4,N}(8N+16k+9) = \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(8N+16k+10) = 2N+121$
- $Q_{4,N}(8N+16k+11) = 3N$
- $Q_{4,N}(8N+16k+12) = 16k+4N+121$
- $Q_{4,N}(8N+16k+13) = \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(8N+16k+14) = 2N+121$
- $Q_{4,N}(8N+16k+15) = 3N$

**Some sporadic values:**

- $Q_{4,N}(9N+22) = 5N+221$
- $Q_{4,N}(9N+23) = \frac{1}{8}N^2 + \frac{23}{4}N$
- $Q_{4,N}(9N+24) = 2N+121$
- $Q_{4,N}(9N+25) = 3N$

- $Q_{4,N}(9N+26) = 5N+341$
- $Q_{4,N}(9N+27) = \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(9N+28) = 2N+121$
- $Q_{4,N}(9N+29) = 3N$
- $Q_{4,N}(9N+30) = 5N+261$
- $Q_{4,N}(9N+31) = \frac{1}{8}N^2 + \frac{23}{4}N$
- $Q_{4,N}(9N+32) = 2N+121$
- $Q_{4,N}(9N+33) = 3N$
- $Q_{4,N}(9N+34) = 5N+501$
- $Q_{4,N}(9N+35) = \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(9N+36) = 2N+121$
- $Q_{4,N}(9N+37) = 4N$
- $Q_{4,N}(9N+38) = 32$
- $Q_{4,N}(9N+39) = \frac{1}{4}N^2 + \frac{27}{2}N$
- $Q_{4,N}(9N+40) = 2N+121$
- $Q_{4,N}(9N+41) = 3N$
- $Q_{4,N}(9N+42) = 5N+501$
- $Q_{4,N}(9N+43) = \frac{1}{8}N^2 + \frac{15}{4}N$
- $Q_{4,N}(9N+44) = 2N+121$
- $Q_{4,N}(9N+45) = 3N$
- $Q_{4,N}(9N+46) = 5N+541$
- $Q_{4,N}(9N+47) = \frac{1}{8}N^2 + \frac{15}{4}N$
- $Q_{4,N}(9N+48) = 2N+121$
- $Q_{4,N}(9N+49) = 3N$
- $Q_{4,N}(9N+50) = 5N+581$

**For indices from $9N+51$ to $10N$:**

- $Q_{4,N}(9N+64k) = 2N+121$
- $Q_{4,N}(9N+64k+1) = 2N$
- $Q_{4,N}(9N+64k+2) = 64$
- $Q_{4,N}(9N+64k+3) = \frac{8}{3}Nk^4 + 12Nk^3 + \frac{1}{8}N^2k^2 + \frac{193}{12}Nk^2 + \frac{1}{8}N^2k + \frac{3}{4}Nk + \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(9N+64k+4) = 2N+121$
- $Q_{4,N}(9N+64k+5) = 2N$
- $Q_{4,N}(9N+64k+6) = 64$
- $Q_{4,N}(9N+64k+7) = \frac{8}{3}Nk^4 + \frac{40}{3}Nk^3 + \frac{1}{8}N^2k^2 + \frac{217}{12}Nk^2 + \frac{1}{8}N^2k + \frac{5}{12}Nk + \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(9N+64k+8) = 2N+121$
- $Q_{4,N}(9N+64k+9) = 2N$
- $Q_{4,N}(9N+64k+10) = 64$
- $Q_{4,N}(9N+64k+11) = \frac{8}{3}Nk^4 + \frac{40}{3}Nk^3 + \frac{1}{8}N^2k^2 + \frac{217}{12}Nk^2 + \frac{1}{4}N^2k + \frac{61}{6}Nk + \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(9N+64k+12) = 2N+121$
- $Q_{4,N}(9N+64k+13) = 2N$

- $Q_{4,N}(9N+64k+14)=64$
- $Q_{4,N}(9N+64k+15)=\frac{8}{3}Nk^4+\frac{44}{3}Nk^3+\frac{1}{8}N^2k^2+\frac{337}{12}Nk^2+\frac{1}{4}N^2k+\frac{113}{6}Nk+\frac{1}{8}N^2+\frac{19}{4}N$
- $Q_{4,N}(9N+64k+16)=2N+121$
- $Q_{4,N}(9N+64k+17)=2N$
- $Q_{4,N}(9N+64k+18)=64$
- $Q_{4,N}(9N+64k+19)=\frac{8}{3}Nk^4+\frac{44}{3}Nk^3+\frac{1}{8}N^2k^2+\frac{313}{12}Nk^2+\frac{1}{4}N^2k+\frac{113}{6}Nk+\frac{1}{8}N^2+\frac{19}{4}N$
- $Q_{4,N}(9N+64k+20)=2N+121$
- $Q_{4,N}(9N+64k+21)=2N$
- $Q_{4,N}(9N+64k+22)=64$
- $Q_{4,N}(9N+64k+23)=\frac{8}{3}Nk^4+16Nk^3+\frac{1}{8}N^2k^2+\frac{349}{12}Nk^2+\frac{1}{4}N^2k+\frac{37}{2}Nk+\frac{1}{8}N^2+\frac{23}{4}N$
- $Q_{4,N}(9N+64k+24)=2N+121$
- $Q_{4,N}(9N+64k+25)=2N$
- $Q_{4,N}(9N+64k+26)=64$
- $Q_{4,N}(9N+64k+27)=\frac{8}{3}Nk^4+16Nk^3+\frac{1}{8}N^2k^2+\frac{349}{12}Nk^2+\frac{1}{4}N^2k+\frac{47}{2}Nk+\frac{1}{8}N^2+\frac{11}{4}N$
- $Q_{4,N}(9N+64k+28)=2N+121$
- $Q_{4,N}(9N+64k+29)=2N$
- $Q_{4,N}(9N+64k+30)=64$
- $Q_{4,N}(9N+64k+31)=\frac{8}{3}Nk^4+\frac{52}{3}Nk^3+\frac{1}{8}N^2k^2+\frac{481}{12}Nk^2+\frac{1}{4}N^2k+\frac{175}{6}Nk+\frac{1}{8}N^2+\frac{23}{4}N$
- $Q_{4,N}(9N+64k+32)=2N+121$
- $Q_{4,N}(9N+64k+33)=2N$
- $Q_{4,N}(9N+64k+34)=64$
- $Q_{4,N}(9N+64k+35)=\frac{8}{3}Nk^4+\frac{52}{3}Nk^3+\frac{1}{8}N^2k^2+\frac{457}{12}Nk^2+\frac{1}{4}N^2k+\frac{163}{6}Nk+\frac{1}{8}N^2+\frac{11}{4}N$
- $Q_{4,N}(9N+64k+36)=2N+121$
- $Q_{4,N}(9N+64k+37)=2N$
- $Q_{4,N}(9N+64k+38)=64$
- $Q_{4,N}(9N+64k+39)=\frac{8}{3}Nk^4+\frac{56}{3}Nk^3+\frac{1}{8}N^2k^2+\frac{505}{12}Nk^2+\frac{1}{4}N^2k+\frac{179}{6}Nk+\frac{1}{4}N^2+\frac{27}{2}N$
- $Q_{4,N}(9N+64k+40)=2N+121$
- $Q_{4,N}(9N+64k+41)=2N$
- $Q_{4,N}(9N+64k+42)=64$
- $Q_{4,N}(9N+64k+43)=\frac{8}{3}Nk^4+\frac{56}{3}Nk^3+\frac{1}{8}N^2k^2+\frac{505}{12}Nk^2+\frac{3}{8}N^2k+\frac{475}{12}Nk+\frac{1}{8}N^2+\frac{15}{4}N$
- $Q_{4,N}(9N+64k+44)=2N+121$
- $Q_{4,N}(9N+64k+45)=2N$
- $Q_{4,N}(9N+64k+46)=64$

- $Q_{4,N}(9N+64k+47) = \frac{8}{3}Nk^4 + 20Nk^3 + \frac{1}{8}N^2k^2 + \frac{649}{12}Nk^2 + \frac{3}{8}N^2k + \frac{237}{4}Nk + \frac{1}{8}N^2 + \frac{15}{4}N$
- $Q_{4,N}(9N+64k+48) = 2N+121$
- $Q_{4,N}(9N+64k+49) = 2N$
- $Q_{4,N}(9N+64k+50) = 64$
- $Q_{4,N}(9N+64k+51) = \frac{8}{3}Nk^4 + 20Nk^3 + \frac{1}{8}N^2k^2 + \frac{625}{12}Nk^2 + \frac{3}{8}N^2k + \frac{229}{4}Nk + \frac{1}{8}N^2 + \frac{15}{4}N$
- $Q_{4,N}(9N+64k+52) = 2N+121$
- $Q_{4,N}(9N+64k+53) = 2N$
- $Q_{4,N}(9N+64k+54) = 64$
- $Q_{4,N}(9N+64k+55) = \frac{8}{3}Nk^4 + \frac{64}{3}Nk^3 + \frac{1}{8}N^2k^2 + \frac{685}{12}Nk^2 + \frac{3}{8}N^2k + \frac{731}{12}Nk + \frac{3}{8}N^2 + \frac{101}{4}N$
- $Q_{4,N}(9N+64k+56) = 2N+121$
- $Q_{4,N}(9N+64k+57) = 2N$
- $Q_{4,N}(9N+64k+58) = 64$
- $Q_{4,N}(9N+64k+59) = \frac{8}{3}Nk^4 + \frac{64}{3}Nk^3 + \frac{1}{8}N^2k^2 + \frac{685}{12}Nk^2 + \frac{3}{8}N^2k + \frac{791}{12}Nk + \frac{3}{8}N^2 + \frac{129}{4}N$
- $Q_{4,N}(9N+64k+60) = 2N+121$
- $Q_{4,N}(9N+64k+61) = 2N$
- $Q_{4,N}(9N+64k+62) = 64$
- $Q_{4,N}(9N+64k+63) = \frac{8}{3}Nk^4 + \frac{68}{3}Nk^3 + \frac{1}{8}N^2k^2 + \frac{841}{12}Nk^2 + \frac{3}{8}N^2k + \frac{1003}{12}Nk + \frac{3}{8}N^2 + \frac{153}{4}N$

**Some sporadic values:**

- $Q_{4,N}(10N+1) = \frac{1}{6291456}N^5 + \frac{59}{786432}N^4 + \frac{4315}{786432}N^3 + \frac{23155}{196608}N^2 + \frac{752603}{131072}N$
- $Q_{4,N}(10N+2) = 2N+121$
- $Q_{4,N}(10N+3) = 2N$
- $Q_{4,N}(10N+4) = 68$
- $Q_{4,N}(10N+5) = \frac{1}{6291456}N^5 + \frac{59}{786432}N^4 + \frac{4315}{786432}N^3 + \frac{23155}{196608}N^2 + \frac{752603}{131072}N$
- $Q_{4,N}(10N+6) = 2N+121$
- $Q_{4,N}(10N+7) = 2N$
- $Q_{4,N}(10N+8) = 64$
- $Q_{4,N}(10N+9) = \frac{1}{6291456}N^5 + \frac{21}{262144}N^4 + \frac{6211}{786432}N^3 + \frac{16981}{65536}N^2 + \frac{845355}{131072}N$
- $Q_{4,N}(10N+10) = 2N+121$
- $Q_{4,N}(10N+11) = 2N$
- $Q_{4,N}(10N+12) = 64$
- $Q_{4,N}(10N+13) = \frac{1}{6291456}N^5 + \frac{67}{786432}N^4 + \frac{8107}{786432}N^3 + \frac{75659}{196608}N^2 + \frac{680059}{131072}N$

- $Q_{4,N}(10N+14) = 2N + 121$
- $Q_{4,N}(10N+15) = 2N$
- $Q_{4,N}(10N+16) = 76$
- $Q_{4,N}(10N+17) = \frac{1}{6291456}N^5 + \frac{21}{262144}N^4 + \frac{4675}{786432}N^3 + \frac{19541}{65536}N^2 + \frac{2671147}{131072}N$
- $Q_{4,N}(10N+18) = 2N + 121$
- $Q_{4,N}(10N+19) = 2N$
- $Q_{4,N}(10N+20) = 88$
- $Q_{4,N}(10N+21) = \frac{1}{6291456}N^5 + \frac{59}{786432}N^4 + \frac{4699}{786432}N^3 + \frac{22771}{196608}N^2 + \frac{752859}{131072}N$
- $Q_{4,N}(10N+22) = 2N + 121$
- $Q_{4,N}(10N+23) = 2N$
- $Q_{4,N}(10N+24) = 2N + 801$
- $Q_{4,N}(10N+25) = \frac{1}{24576}N^4 + \frac{23}{4096}N^3 - \frac{205}{6144}N^2 + \frac{7213}{1024}N$
- $Q_{4,N}(10N+26) = 2N + 121$
- $Q_{4,N}(10N+27) = 2N$
- $Q_{4,N}(10N+28) = 108$
- $Q_{4,N}(10N+29) = \frac{1}{6291456}N^5 + \frac{17}{262144}N^4 + \frac{3595}{786432}N^3 - \frac{4663}{65536}N^2 - \frac{1163461}{131072}N$
- $Q_{4,N}(10N+30) = 2N + 121$
- $Q_{4,N}(10N+31) = 2N$
- $Q_{4,N}(10N+32) = 2N + 1361$
- $Q_{4,N}(10N+33) = \frac{1}{24576}N^4 - \frac{1}{4096}N^3 + \frac{131}{6144}N^2 + \frac{11221}{1024}N$
- $Q_{4,N}(10N+34) = 2N + 121$
- $Q_{4,N}(10N+35) = 2N$
- $Q_{4,N}(10N+36) = 6N + 921$
- $Q_{4,N}(10N+37) = \frac{1}{4}N^2 - \frac{5}{2}N$
- $Q_{4,N}(10N+38) = 2N + 121$
- $Q_{4,N}(10N+39) = 2N$
- $Q_{4,N}(10N+40) = 2N + 1561$
- $Q_{4,N}(10N+41) = \frac{1}{24576}N^4 - \frac{5}{4096}N^3 + \frac{1883}{6144}N^2 - \frac{12911}{1024}N$
- $Q_{4,N}(10N+42) = 2N + 121$
- $Q_{4,N}(10N+43) = 2N$
- $Q_{4,N}(10N+44) = 6N + 1321$
- $Q_{4,N}(10N+45) = \frac{1}{4}N^2 - \frac{13}{2}N$
- $Q_{4,N}(10N+46) = 2N + 121$
- $Q_{4,N}(10N+47) = 3N$
- $Q_{4,N}(10N+48) = 6N + 561$
- $Q_{4,N}(10N+49) = \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(10N+50) = 2N + 121$
- $Q_{4,N}(10N+51) = 2N$

- $Q_{4,N}(10N+52) = 6N+1401$
- $Q_{4,N}(10N+53) = \frac{1}{4}N^2 - \frac{9}{2}N$
- $Q_{4,N}(10N+54) = 2N+121$
- $Q_{4,N}(10N+55) = 3N$
- $Q_{4,N}(10N+56) = 6N+641$
- $Q_{4,N}(10N+57) = \frac{1}{8}N^2 + \frac{11}{4}N$

**For indices from $10N+58$ to $11N+17$:**

- $Q_{4,N}(10N+32k) = 32k+6N$
- $Q_{4,N}(10N+32k+1) = \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(10N+32k+2) = 2N+121$
- $Q_{4,N}(10N+32k+3) = 3N$
- $Q_{4,N}(10N+32k+4) = 32k+6N+121$
- $Q_{4,N}(10N+32k+5) = \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(10N+32k+6) = 2N+121$
- $Q_{4,N}(10N+32k+7) = 3N$
- $Q_{4,N}(10N+32k+8) = 32k+6N+81$
- $Q_{4,N}(10N+32k+9) = \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(10N+32k+10) = 2N+121$
- $Q_{4,N}(10N+32k+11) = 3N$
- $Q_{4,N}(10N+32k+12) = 32k+6N+121$
- $Q_{4,N}(10N+32k+13) = \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(10N+32k+14) = 2N+121$
- $Q_{4,N}(10N+32k+15) = 3N$
- $Q_{4,N}(10N+32k+16) = 32k+6N+161$
- $Q_{4,N}(10N+32k+17) = \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(10N+32k+18) = 2N+121$
- $Q_{4,N}(10N+32k+19) = 3N$
- $Q_{4,N}(10N+32k+20) = 32k+6N+201$
- $Q_{4,N}(10N+32k+21) = \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(10N+32k+22) = 2N+121$
- $Q_{4,N}(10N+32k+23) = 3N$
- $Q_{4,N}(10N+32k+24) = 32k+6N+241$
- $Q_{4,N}(10N+32k+25) = \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(10N+32k+26) = 2N+121$
- $Q_{4,N}(10N+32k+27) = 3N$
- $Q_{4,N}(10N+32k+28) = 32k+6N+281$
- $Q_{4,N}(10N+32k+29) = \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(10N+32k+30) = 2N+121$
- $Q_{4,N}(10N+32k+31) = 3N$

**Some sporadic values:**

- $Q_{4,N}(11N+18)=7N+261$
- $Q_{4,N}(11N+19)=\frac{1}{8}N^2+\frac{11}{4}N$
- $Q_{4,N}(11N+20)=2N+121$
- $Q_{4,N}(11N+21)=3N$
- $Q_{4,N}(11N+22)=7N+221$
- $Q_{4,N}(11N+23)=\frac{1}{8}N^2+\frac{19}{4}N$
- $Q_{4,N}(11N+24)=2N+121$
- $Q_{4,N}(11N+25)=3N$
- $Q_{4,N}(11N+26)=7N+261$
- $Q_{4,N}(11N+27)=\frac{1}{8}N^2+\frac{19}{4}N$
- $Q_{4,N}(11N+28)=2N+121$
- $Q_{4,N}(11N+29)=3N$
- $Q_{4,N}(11N+30)=7N+301$
- $Q_{4,N}(11N+31)=\frac{1}{8}N^2+\frac{19}{4}N$
- $Q_{4,N}(11N+32)=2N+121$
- $Q_{4,N}(11N+33)=3N$
- $Q_{4,N}(11N+34)=7N+341$
- $Q_{4,N}(11N+35)=\frac{1}{8}N^2+\frac{23}{4}N$
- $Q_{4,N}(11N+36)=2N+121$
- $Q_{4,N}(11N+37)=3N$
- $Q_{4,N}(11N+38)=7N+461$
- $Q_{4,N}(11N+39)=\frac{1}{8}N^2+\frac{11}{4}N$
- $Q_{4,N}(11N+40)=2N+121$
- $Q_{4,N}(11N+41)=3N$
- $Q_{4,N}(11N+42)=7N+381$
- $Q_{4,N}(11N+43)=\frac{1}{8}N^2+\frac{23}{4}N$
- $Q_{4,N}(11N+44)=2N+121$
- $Q_{4,N}(11N+45)=3N$
- $Q_{4,N}(11N+46)=7N+621$
- $Q_{4,N}(11N+47)=\frac{1}{8}N^2+\frac{11}{4}N$
- $Q_{4,N}(11N+48)=2N+121$
- $Q_{4,N}(11N+49)=4N$
- $Q_{4,N}(11N+50)=64$
- $Q_{4,N}(11N+51)=\frac{3}{8}N^2+\frac{109}{4}N$
- $Q_{4,N}(11N+52)=2N+121$
- $Q_{4,N}(11N+53)=3N$
- $Q_{4,N}(11N+54)=7N+621$
- $Q_{4,N}(11N+55)=\frac{1}{8}N^2+\frac{15}{4}N$
- $Q_{4,N}(11N+56)=2N+121$
- $Q_{4,N}(11N+57)=3N$
- $Q_{4,N}(11N+58)=7N+661$
- $Q_{4,N}(11N+59)=\frac{1}{8}N^2+\frac{15}{4}N$
- $Q_{4,N}(11N+60)=2N+121$
- $Q_{4,N}(11N+61)=3N$
- $Q_{4,N}(11N+62)=7N+701$

**For indices from $11N+63$ to $12N$:**

- $Q_{4,N}(11N+128k)=2N+121$
- $Q_{4,N}(11N+128k+1)=2N$
- $Q_{4,N}(11N+128k+2)=128$
- $Q_{4,N}(11N+128k+3)=\frac{128}{15}Nk^5+\frac{136}{3}Nk^4+\frac{1}{6}N^2k^3+\frac{251}{3}Nk^3+\frac{3}{8}N^2k^2+\frac{683}{12}Nk^2+\frac{1}{3}N^2k+\frac{74}{5}Nk+\frac{1}{8}N^2+\frac{11}{4}N$
- $Q_{4,N}(11N+128k+4)=2N+121$
- $Q_{4,N}(11N+128k+5)=2N$
- $Q_{4,N}(11N+128k+6)=128$
- $Q_{4,N}(11N+128k+7)=\frac{128}{15}Nk^5+48Nk^4+\frac{1}{6}N^2k^3+\frac{275}{3}Nk^3+\frac{3}{8}N^2k^2+\frac{253}{4}Nk^2+\frac{1}{3}N^2k+\frac{69}{5}Nk+\frac{1}{8}N^2+\frac{19}{4}N$
- $Q_{4,N}(11N+128k+8)=2N+121$
- $Q_{4,N}(11N+128k+9)=2N$
- $Q_{4,N}(11N+128k+10)=128$
- $Q_{4,N}(11N+128k+11)=\frac{128}{15}Nk^5+48Nk^4+\frac{1}{6}N^2k^3+\frac{275}{3}Nk^3+\frac{1}{2}N^2k^2+73Nk^2+\frac{11}{24}N^2k+\frac{511}{20}Nk+\frac{1}{8}N^2+\frac{19}{4}N$
- $Q_{4,N}(11N+128k+12)=2N+121$
- $Q_{4,N}(11N+128k+13)=2N$
- $Q_{4,N}(11N+128k+14)=128$
- $Q_{4,N}(11N+128k+15)=\frac{128}{15}Nk^5+\frac{152}{3}Nk^4+\frac{1}{6}N^2k^3+\frac{331}{3}Nk^3+\frac{1}{2}N^2k^2+\frac{313}{3}Nk^2+\frac{11}{24}N^2k+\frac{2453}{60}Nk+\frac{1}{8}N^2+\frac{19}{4}N$
- $Q_{4,N}(11N+128k+16)=2N+121$
- $Q_{4,N}(11N+128k+17)=2N$
- $Q_{4,N}(11N+128k+18)=128$
- $Q_{4,N}(11N+128k+19)=\frac{128}{15}Nk^5+\frac{152}{3}Nk^4+\frac{1}{6}N^2k^3+\frac{323}{3}Nk^3+\frac{1}{2}N^2k^2+\frac{301}{3}Nk^2+\frac{11}{24}N^2k+\frac{791}{20}Nk+\frac{1}{8}N^2+\frac{11}{4}N$
- $Q_{4,N}(11N+128k+20)=2N+121$
- $Q_{4,N}(11N+128k+21)=2N$
- $Q_{4,N}(11N+128k+22)=128$
- $Q_{4,N}(11N+128k+23)=\frac{128}{15}Nk^5+\frac{160}{3}Nk^4+\frac{1}{6}N^2k^3+117Nk^3+\frac{1}{2}N^2k^2+\frac{326}{3}Nk^2+\frac{11}{24}N^2k+\frac{2533}{60}Nk+\frac{1}{8}N^2+\frac{19}{4}N$
- $Q_{4,N}(11N+128k+24)=2N+121$
- $Q_{4,N}(11N+128k+25)=2N$
- $Q_{4,N}(11N+128k+26)=128$
- $Q_{4,N}(11N+128k+27)=\frac{128}{15}Nk^5+\frac{160}{3}Nk^4+\frac{1}{6}N^2k^3+117Nk^3+\frac{1}{2}N^2k^2+\frac{341}{3}Nk^2+\frac{11}{24}N^2k+\frac{2653}{60}Nk+\frac{1}{8}N^2+\frac{19}{4}N$

- $Q_{4,N}(11N + 128k + 28) = 2N + 121$
- $Q_{4,N}(11N + 128k + 29) = 2N$
- $Q_{4,N}(11N + 128k + 30) = 128$
- $Q_{4,N}(11N + 128k + 31) = \frac{128}{15}Nk^5 + 56Nk^4 + \frac{1}{6}N^2k^3 + 137Nk^3 + \frac{1}{2}N^2k^2 + 144Nk^2 + \frac{11}{24}N^2k + \frac{3613}{60}Nk + \frac{1}{8}N^2 + \frac{19}{4}N$
- $Q_{4,N}(11N + 128k + 32) = 2N + 121$
- $Q_{4,N}(11N + 128k + 33) = 2N$
- $Q_{4,N}(11N + 128k + 34) = 128$
- $Q_{4,N}(11N + 128k + 35) = \frac{128}{15}Nk^5 + 56Nk^4 + \frac{1}{6}N^2k^3 + \frac{403}{3}Nk^3 + \frac{1}{2}N^2k^2 + 138Nk^2 + \frac{11}{24}N^2k + \frac{3233}{60}Nk + \frac{1}{8}N^2 + \frac{23}{4}N$
- $Q_{4,N}(11N + 128k + 36) = 2N + 121$
- $Q_{4,N}(11N + 128k + 37) = 2N$
- $Q_{4,N}(11N + 128k + 38) = 128$
- $Q_{4,N}(11N + 128k + 39) = \frac{128}{15}Nk^5 + \frac{176}{3}Nk^4 + \frac{1}{6}N^2k^3 + 145Nk^3 + \frac{1}{2}N^2k^2 + \frac{454}{3}Nk^2 + \frac{7}{12}N^2k + \frac{2099}{30}Nk + \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(11N + 128k + 40) = 2N + 121$
- $Q_{4,N}(11N + 128k + 41) = 2N$
- $Q_{4,N}(11N + 128k + 42) = 128$
- $Q_{4,N}(11N + 128k + 43) = \frac{128}{15}Nk^5 + \frac{176}{3}Nk^4 + \frac{1}{6}N^2k^3 + 145Nk^3 + \frac{5}{8}N^2k^2 + \frac{1933}{12}Nk^2 + \frac{7}{12}N^2k + \frac{2099}{30}Nk + \frac{1}{8}N^2 + \frac{23}{4}N$
- $Q_{4,N}(11N + 128k + 44) = 2N + 121$
- $Q_{4,N}(11N + 128k + 45) = 2N$
- $Q_{4,N}(11N + 128k + 46) = 128$
- $Q_{4,N}(11N + 128k + 47) = \frac{128}{15}Nk^5 + \frac{184}{3}Nk^4 + \frac{1}{6}N^2k^3 + \frac{499}{3}Nk^3 + \frac{5}{8}N^2k^2 + \frac{2489}{12}Nk^2 + \frac{7}{12}N^2k + \frac{2929}{30}Nk + \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(11N + 128k + 48) = 2N + 121$
- $Q_{4,N}(11N + 128k + 49) = 2N$
- $Q_{4,N}(11N + 128k + 50) = 128$
- $Q_{4,N}(11N + 128k + 51) = \frac{128}{15}Nk^5 + \frac{184}{3}Nk^4 + \frac{1}{6}N^2k^3 + \frac{491}{3}Nk^3 + \frac{5}{8}N^2k^2 + \frac{2417}{12}Nk^2 + \frac{7}{12}N^2k + \frac{943}{10}Nk + \frac{3}{8}N^2 + \frac{109}{4}N$
- $Q_{4,N}(11N + 128k + 52) = 2N + 121$
- $Q_{4,N}(11N + 128k + 53) = 2N$
- $Q_{4,N}(11N + 128k + 54) = 128$
- $Q_{4,N}(11N + 128k + 55) = \frac{128}{15}Nk^5 + 64Nk^4 + \frac{1}{6}N^2k^3 + \frac{527}{3}Nk^3 + \frac{5}{8}N^2k^2 + \frac{871}{4}Nk^2 + \frac{5}{6}N^2k + \frac{614}{5}Nk + \frac{1}{8}N^2 + \frac{15}{4}N$
- $Q_{4,N}(11N + 128k + 56) = 2N + 121$

- $Q_{4,N}(11N + 128k + 57) = 2N$
- $Q_{4,N}(11N + 128k + 58) = 128$
- $Q_{4,N}(11N + 128k + 59) = \frac{128}{15}Nk^5 + 64Nk^4 + \frac{1}{6}N^2k^3 + \frac{527}{3}Nk^3 + \frac{5}{8}N^2k^2 + \frac{891}{4}Nk^2 + \frac{5}{6}N^2k + \frac{674}{5}Nk + \frac{1}{8}N^2 + \frac{15}{4}N$
- $Q_{4,N}(11N + 128k + 60) = 2N + 121$
- $Q_{4,N}(11N + 128k + 61) = 2N$
- $Q_{4,N}(11N + 128k + 62) = 128$
- $Q_{4,N}(11N + 128k + 63) = \frac{128}{15}Nk^5 + \frac{200}{3}Nk^4 + \frac{1}{6}N^2k^3 + \frac{595}{3}Nk^3 + \frac{5}{8}N^2k^2 + \frac{3229}{12}Nk^2 + \frac{5}{6}N^2k + \frac{2507}{15}Nk + \frac{1}{8}N^2 + \frac{15}{4}N$
- $Q_{4,N}(11N + 128k + 64) = 2N + 121$
- $Q_{4,N}(11N + 128k + 65) = 2N$
- $Q_{4,N}(11N + 128k + 66) = 128$
- $Q_{4,N}(11N + 128k + 67) = \frac{128}{15}Nk^5 + \frac{200}{3}Nk^4 + \frac{1}{6}N^2k^3 + \frac{587}{3}Nk^3 + \frac{5}{8}N^2k^2 + \frac{3133}{12}Nk^2 + \frac{5}{6}N^2k + \frac{799}{5}Nk + \frac{1}{2}N^2 + 39N$
- $Q_{4,N}(11N + 128k + 68) = 2N + 121$
- $Q_{4,N}(11N + 128k + 69) = 2N$
- $Q_{4,N}(11N + 128k + 70) = 128$
- $Q_{4,N}(11N + 128k + 71) = \frac{128}{15}Nk^5 + \frac{208}{3}Nk^4 + \frac{1}{6}N^2k^3 + 209Nk^3 + \frac{5}{8}N^2k^2 + \frac{3401}{12}Nk^2 + \frac{5}{6}N^2k + \frac{2587}{15}Nk + \frac{1}{2}N^2 + 42N$
- $Q_{4,N}(11N + 128k + 72) = 2N + 121$
- $Q_{4,N}(11N + 128k + 73) = 2N$
- $Q_{4,N}(11N + 128k + 74) = 128$
- $Q_{4,N}(11N + 128k + 75) = \frac{128}{15}Nk^5 + \frac{208}{3}Nk^4 + \frac{1}{6}N^2k^3 + 209Nk^3 + \frac{3}{4}N^2k^2 + \frac{1759}{6}Nk^2 + \frac{13}{12}N^2k + \frac{5819}{30}Nk + \frac{5}{8}N^2 + \frac{215}{4}N$
- $Q_{4,N}(11N + 128k + 76) = 2N + 121$
- $Q_{4,N}(11N + 128k + 77) = 2N$
- $Q_{4,N}(11N + 128k + 78) = 128$
- $Q_{4,N}(11N + 128k + 79) = \frac{128}{15}Nk^5 + 72Nk^4 + \frac{1}{6}N^2k^3 + 233Nk^3 + \frac{3}{4}N^2k^2 + \frac{713}{2}Nk^2 + \frac{13}{12}N^2k + \frac{7679}{30}Nk + \frac{5}{8}N^2 + \frac{295}{4}N$
- $Q_{4,N}(11N + 128k + 80) = 2N + 121$
- $Q_{4,N}(11N + 128k + 81) = 2N$
- $Q_{4,N}(11N + 128k + 82) = 128$
- $Q_{4,N}(11N + 128k + 83) = \frac{128}{15}Nk^5 + 72Nk^4 + \frac{1}{6}N^2k^3 + \frac{691}{3}Nk^3 + \frac{3}{4}N^2k^2 + \frac{697}{2}Nk^2 + \frac{13}{12}N^2k + \frac{7459}{30}Nk + \frac{5}{8}N^2 + \frac{287}{4}N$
- $Q_{4,N}(11N + 128k + 84) = 2N + 121$
- $Q_{4,N}(11N + 128k + 85) = 2N$

- $Q_{4,N}(11N + 128k + 86) = 128$
- $Q_{4,N}(11N + 128k + 87) = \frac{128}{15}Nk^5 + \frac{224}{3}Nk^4 + \frac{1}{6}N^2k^3 + 245Nk^3 + \frac{3}{4}N^2k^2 + \frac{2249}{6}Nk^2 + \frac{13}{12}N^2k + \frac{8039}{30}Nk + \frac{5}{8}N^2 + \frac{307}{4}N$
- $Q_{4,N}(11N + 128k + 88) = 2N + 121$
- $Q_{4,N}(11N + 128k + 89) = 2N$
- $Q_{4,N}(11N + 128k + 90) = 128$
- $Q_{4,N}(11N + 128k + 91) = \frac{128}{15}Nk^5 + \frac{224}{3}Nk^4 + \frac{1}{6}N^2k^3 + 245Nk^3 + \frac{3}{4}N^2k^2 + \frac{2279}{6}Nk^2 + \frac{13}{12}N^2k + \frac{8249}{30}Nk + \frac{5}{8}N^2 + \frac{315}{4}N$
- $Q_{4,N}(11N + 128k + 92) = 2N + 121$
- $Q_{4,N}(11N + 128k + 93) = 2N$
- $Q_{4,N}(11N + 128k + 94) = 128$
- $Q_{4,N}(11N + 128k + 95) = \frac{128}{15}Nk^5 + \frac{232}{3}Nk^4 + \frac{1}{6}N^2k^3 + \frac{811}{3}Nk^3 + \frac{3}{4}N^2k^2 + \frac{2665}{6}Nk^2 + \frac{13}{12}N^2k + \frac{10129}{30}Nk + \frac{5}{8}N^2 + \frac{399}{4}N$
- $Q_{4,N}(11N + 128k + 96) = 2N + 121$
- $Q_{4,N}(11N + 128k + 97) = 2N$
- $Q_{4,N}(11N + 128k + 98) = 128$
- $Q_{4,N}(11N + 128k + 99) = \frac{128}{15}Nk^5 + \frac{232}{3}Nk^4 + \frac{1}{6}N^2k^3 + \frac{803}{3}Nk^3 + \frac{3}{4}N^2k^2 + \frac{2605}{6}Nk^2 + \frac{13}{12}N^2k + \frac{3233}{10}Nk + \frac{5}{8}N^2 + \frac{363}{4}N$
- $Q_{4,N}(11N + 128k + 100) = 2N + 121$
- $Q_{4,N}(11N + 128k + 101) = 2N$
- $Q_{4,N}(11N + 128k + 102) = 128$
- $Q_{4,N}(11N + 128k + 103) = \frac{128}{15}Nk^5 + 80Nk^4 + \frac{1}{6}N^2k^3 + \frac{851}{3}Nk^3 + \frac{3}{4}N^2k^2 + \frac{935}{2}Nk^2 + \frac{29}{24}N^2k + \frac{7241}{20}Nk + \frac{3}{4}N^2 + \frac{223}{2}N$
- $Q_{4,N}(11N + 128k + 104) = 2N + 121$
- $Q_{4,N}(11N + 128k + 105) = 2N$
- $Q_{4,N}(11N + 128k + 106) = 128$
- $Q_{4,N}(11N + 128k + 107) = \frac{128}{15}Nk^5 + 80Nk^4 + \frac{1}{6}N^2k^3 + \frac{851}{3}Nk^3 + \frac{7}{8}N^2k^2 + \frac{1909}{4}Nk^2 + \frac{4}{3}N^2k + \frac{1859}{5}Nk + \frac{3}{4}N^2 + \frac{223}{2}N$
- $Q_{4,N}(11N + 128k + 108) = 2N + 121$
- $Q_{4,N}(11N + 128k + 109) = 2N$
- $Q_{4,N}(11N + 128k + 110) = 128$
- $Q_{4,N}(11N + 128k + 111) = \frac{128}{15}Nk^5 + \frac{248}{3}Nk^4 + \frac{1}{6}N^2k^3 + \frac{931}{3}Nk^3 + \frac{7}{8}N^2k^2 + \frac{6715}{12}Nk^2 + \frac{4}{3}N^2k + \frac{6947}{15}Nk + \frac{3}{4}N^2 + \frac{289}{2}N$
- $Q_{4,N}(11N + 128k + 112) = 2N + 121$
- $Q_{4,N}(11N + 128k + 113) = 2N$
- $Q_{4,N}(11N + 128k + 114) = 128$

- $Q_{4,N}(11N+128k+115) = \frac{128}{15}Nk^5 + \frac{248}{3}Nk^4 + \frac{1}{6}N^2k^3 + \frac{923}{3}Nk^3 + \frac{7}{8}N^2k^2 + \frac{6595}{12}Nk^2 + \frac{4}{3}N^2k + \frac{2259}{5}Nk + \frac{3}{4}N^2 + \frac{281}{2}N$
- $Q_{4,N}(11N+128k+116) = 2N+121$
- $Q_{4,N}(11N+128k+117) = 2N$
- $Q_{4,N}(11N+128k+118) = 128$
- $Q_{4,N}(11N+128k+119) = \frac{128}{15}Nk^5 + \frac{256}{3}Nk^4 + \frac{1}{6}N^2k^3 + 325Nk^3 + \frac{7}{8}N^2k^2 + \frac{7055}{12}Nk^2 + \frac{19}{12}N^2k + \frac{15209}{30}Nk + N^2 + 172N$
- $Q_{4,N}(11N+128k+120) = 2N+121$
- $Q_{4,N}(11N+128k+121) = 2N$
- $Q_{4,N}(11N+128k+122) = 128$
- $Q_{4,N}(11N+128k+123) = \frac{128}{15}Nk^5 + \frac{256}{3}Nk^4 + \frac{1}{6}N^2k^3 + 325Nk^3 + \frac{7}{8}N^2k^2 + \frac{7115}{12}Nk^2 + \frac{19}{12}N^2k + \frac{15719}{30}Nk + N^2 + 184N$
- $Q_{4,N}(11N+128k+124) = 2N+121$
- $Q_{4,N}(11N+128k+125) = 2N$
- $Q_{4,N}(11N+128k+126) = 128$
- $Q_{4,N}(11N+128k+127) = \frac{128}{15}Nk^5 + 88Nk^4 + \frac{1}{6}N^2k^3 + 353Nk^3 + \frac{7}{8}N^2k^2 + \frac{2709}{4}Nk^2 + \frac{19}{12}N^2k + \frac{18629}{30}Nk + N^2 + 222N$

**Some sporadic values:**

- $Q_{4,N}(12N+1) = \frac{1}{4026531840}N^6 + \frac{33}{134217728}N^5 + \frac{6137}{100663296}N^4 + \frac{96505}{16777216}N^3 + \frac{18629863}{83886080}N^2 + \frac{29631325}{8388608}N$
- $Q_{4,N}(12N+2) = 2N+121$
- $Q_{4,N}(12N+3) = 2N$
- $Q_{4,N}(12N+4) = 132$
- $Q_{4,N}(12N+5) = \frac{1}{4026531840}N^6 + \frac{33}{134217728}N^5 + \frac{6137}{100663296}N^4 + \frac{96505}{16777216}N^3 + \frac{18629863}{83886080}N^2 + \frac{29631325}{8388608}N$
- $Q_{4,N}(12N+6) = 2N+121$
- $Q_{4,N}(12N+7) = 2N$
- $Q_{4,N}(12N+8) = 128$
- $Q_{4,N}(12N+9) = \frac{1}{4026531840}N^6 + \frac{103}{402653184}N^5 + \frac{2427}{33554432}N^4 + \frac{385399}{50331648}N^3 + \frac{75564389}{251658240}N^2 + \frac{53420809}{8388608}N$
- $Q_{4,N}(12N+10) = 2N+121$
- $Q_{4,N}(12N+11) = 2N$
- $Q_{4,N}(12N+12) = 128$
- $Q_{4,N}(12N+13) = \frac{1}{4026531840}N^6 + \frac{107}{402653184}N^5 + \frac{2723}{33554432}N^4 + \frac{478979}{50331648}N^3 + \frac{103812629}{251658240}N^2 + \frac{43086005}{8388608}N$
- $Q_{4,N}(12N+14) = 2N+121$
- $Q_{4,N}(12N+15) = 2N$
- $Q_{4,N}(12N+16) = 144$

- $Q_{4,N}(12N+17) = \frac{1}{4026531840}N^6 + \frac{33}{134217728}N^5 + \frac{6649}{100663296}N^4 + \frac{104185}{16777216}N^3 + \frac{101288629}{251658240}N^2 + \frac{152418141}{8388608}N$
- $Q_{4,N}(12N+18) = 2N+121$
- $Q_{4,N}(12N+19) = 2N$
- $Q_{4,N}(12N+20) = 152$
- $Q_{4,N}(12N+21) = \frac{1}{4026531840}N^6 + \frac{33}{134217728}N^5 + \frac{6265}{100663296}N^4 + \frac{100473}{16777216}N^3 + \frac{58269109}{251658240}N^2 + \frac{46241885}{8388608}N$
- $Q_{4,N}(12N+22) = 2N+121$
- $Q_{4,N}(12N+23) = 2N$
- $Q_{4,N}(12N+24) = 2N+1441$
- $Q_{4,N}(12N+25) = \frac{1}{6291456}N^5 + \frac{31}{786432}N^4 + \frac{1795}{786432}N^3 + \frac{57503}{196608}N^2 - \frac{77909}{131072}N$
- $Q_{4,N}(12N+26) = 2N+121$
- $Q_{4,N}(12N+27) = 2N$
- $Q_{4,N}(12N+28) = 184$
- $Q_{4,N}(12N+29) = \frac{1}{4026531840}N^6 + \frac{29}{134217728}N^5 + \frac{4241}{100663296}N^4 + \frac{108653}{16777216}N^3 + \frac{5954103}{83886080}N^2 - \frac{118851495}{8388608}N$
- $Q_{4,N}(12N+30) = 2N+121$
- $Q_{4,N}(12N+31) = 2N$
- $Q_{4,N}(12N+32) = 2N+2241$
- $Q_{4,N}(12N+33) = \frac{1}{6291456}N^5 - \frac{5}{786432}N^4 + \frac{2395}{786432}N^3 - \frac{13645}{196608}N^2 + \frac{6832347}{131072}N$
- $Q_{4,N}(12N+34) = 2N+121$
- $Q_{4,N}(12N+35) = 2N$
- $Q_{4,N}(12N+36) = 8N+1721$
- $Q_{4,N}(12N+37) = \frac{3}{8}N^2 - \frac{71}{4}N$
- $Q_{4,N}(12N+38) = 2N+121$
- $Q_{4,N}(12N+39) = 2N$
- $Q_{4,N}(12N+40) = 2N+2641$
- $Q_{4,N}(12N+41) = \frac{1}{6291456}N^5 - \frac{7}{262144}N^4 + \frac{6715}{786432}N^3 - \frac{41471}{65536}N^2 + \frac{8943259}{131072}N$
- $Q_{4,N}(12N+42) = 2N+121$
- $Q_{4,N}(12N+43) = 2N$
- $Q_{4,N}(12N+44) = 8N+2681$
- $Q_{4,N}(12N+45) = \frac{3}{8}N^2 - \frac{131}{4}N$
- $Q_{4,N}(12N+46) = 2N+121$
- $Q_{4,N}(12N+47) = 2N$
- $Q_{4,N}(12N+48) = 12N+1481$
- $Q_{4,N}(12N+49) = \frac{1}{8}N^2 + \frac{11}{4}N$
- $Q_{4,N}(12N+50) = 0$

# Appendix F

## List of OEIS Sequences Referenced

This dissertation mentions the following sequences in the OEIS [31]:

# Appendix G

## List of Supplemental Computer Content

This appendix alphabetically lists all of the Maple programs and other external text files that are referenced in this thesis.

| File | Type | Page(s) |
|---|---|---|
| Hof1thruN.txt | Output | 115 |
| hof_small_periods.txt | Output | 57 |
| N2N2_17_2_Sporadic.txt | Output | 156 |
| N2N2_41_2_Sporadic.txt | Output | 157 |
| N2N2_mod24_1.txt | Output | 151 |
| N2N2_mod24_9.txt | Output | 151 |
| N2N2_mod24_13.txt | Output | 151 |
| N2N2_mod24_21.txt | Output | 151 |
| N4N4 | Directory | 163 |
| N4N4_explore.txt | Maple | 151, 154, 155 157 |
| nicehof.txt | Maple | 19, 172 |
| nonstdhof.txt | Maple | 96, 120, 150, 150, 163, 171, 172 |
| ProveTriHof1thruN.txt | Maple | 126, 127, 130, 131, 134 |
| RSTsearch.txt | Maple | 88 |
| slowsearch.txt | Maple | 145, 171 |
| slowseqs.txt | Output | 145, 171 |
| TriHof193Sporadic.txt | Output | 126 |
| TriHof3442Sporadic.txt | Output | 127 |
| TriHof19395final.txt | Output | 130 |
| TriHof19395Lfinal.txt | Output | 130 |

| **File** | **Type** | **Page(s)** |
|---|---|---|
| `TriHof20830final.txt` | Output | 134 |
| `TriHof20830Lfinal.txt` | Output | 134 |
| `TriHof20830Lmid.txt` | Output | 134 |
| `TriHof20830mid.txt` | Output | 134 |
| `TriHof27298final.txt` | Output | 131 |
| `TriHof27298Lfinal.txt` | Output | 131 |
| `TriHof1thruN.txt` | Output | 122 |
| `trihofform.txt` | Output | 120 |

# Bibliography

[1] J.P. Allouche and J. Shallit, *A variant of Hofstadter's sequence and finite automata*, arXiv preprint arXiv:1103.1133 (2011).

[2] David H Bailey, Jonathan Borwein, Neil Calkin, Roland Girgensohn, Russell Luke, and Victor Moll, *Experimental mathematics in action*, Vol. 174, AK Peters, Natick, MA, 2007.

[3] B. Balamohan, A. Kuznetsov, and Stephen Tanny, *On the behavior of a variant of Hofstadter's Q-sequence*, J. Integer Seq. **10** (2007), 29.

[4] Béla Bollobás, *Modern Graph Theory*, Vol. 184, Springer Science & Business Media, New York, 1998.

[5] Marcel Celaya and Frank Ruskey, *An undecidable nested recurrence relation*, Conference on computability in europe, 2012, pp. 107–117.

[6] Marc Chamberland, *An update on the 3x+ 1 problem*, Butl. Soc. Catalana Mat. **18** (2003), no. 1, 19–45.

[7] B. W. Conolly, *Meta-Fibonacci sequences*, *Fibonacci & Lucas Numbers, and the Golden Section*, 1989, pp. 127–138.

[8] Chris Deugau and Frank Ruskey, *Complete k-ary trees and generalized meta-Fibonacci sequences*, Fourth colloquium on mathematics and computer science algorithms, trees, combinatorics and probabilities, 2006, pp. 203–214.

[9] Alejandro Erickson, Abraham Isgur, Bradley W. Jackson, Frank Ruskey, and Stephen M. Tanny, *Nested recurrence relations with Conolly-like solutions*, SIAM J. Discrete Math. **26** (2012), no. 1, 206–238.

[10] Nathan Fox, *Finding linear-recurrent solutions to Hofstadter-like recurrences using symbolic computation*, arXiv preprint arXiv:1609.06342 (2016).

[11] ———, *Linear recurrent subsequences of generalized meta-Fibonacci sequences*, J. Difference Equ. Appl. (2016).

[12] ———, *Quasipolynomial solutions to the Hofstadter Q-recurrence*, Integers **16** (2016), A68.

[13] ———, *A slow relative of Hofstadter's Q-sequence*, arXiv preprint arXiv:1611.08244 (2016).

[14] Kurt Gödel, *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I*, Monatshefte für mathematik und physik **38** (1931), no. 1, 173–198.

[15] S.W. Golomb, *Discrete chaos: Sequences satisfying "Strange" recursions* (1991).

[16] Richard K Guy, *Unsolved problems in number theory*, Problem E31 (1994).

[17] Kurt Hensel, *Über eine neue Begründung der Theorie der algebraischen Zahlen.*, Jahresber. Dtsch. Math.-Ver. **6** (1897), 83–88.

[18] Douglas Hofstadter, *Gödel, Escher, Bach: an Eternal Golden Braid*, Basic Books, New York, 1979.

[19] ______, *Curious patterns and nonpatterns in a family of meta-Fibonacci recursions*, 2014. Talk given at Rutgers Experimental Mathematics Seminar.

[20] A. Isgur, R. Lech, S. Moore, S. Tanny, Y. Verberne, and Y. Zhang, *Constructing new families of nested recursions with slow solutions*, SIAM J. Discrete Math. **30** (2016), no. 2, 1128–1147.

[21] Abraham Isgur, *Solving nested recursions with trees*, Ph.D. Thesis, 2012.

[22] Abraham Isgur, David Reiss, and Stephen Tanny, *Trees and meta-Fibonacci sequences*, Electron. J. Combin. **16** (2009), no. R129, 1.

[23] Brad Jackson and Frank Ruskey, *Meta-Fibonacci sequences, binary trees and extremal compact codes*, Electron. J. Combin. **13** (2006), no. 1, R26.

[24] Richard M. Karp, *Reducibility among combinatorial problems*, Complexity of computer computations, 1972, pp. 85–103.

[25] Richard Kaye, *Models of peano arithmetic*, Clarendon Press, Oxford University Press, Oxford and New York, 1991.

[26] Dénes König, *Über eine Schlussweise aus dem Endlichen ins Unendliche*, Acta Litt. Ac. Sci. Hung. Fran. Josep **3** (1927), 121–130.

[27] Jeffrey C Lagarias, *The 3x+ 1 problem and its generalizations*, Amer. Math. Monthly **92** (1985), no. 1, 3–23.

[28] Colin L Mallows, *Conway's challenge sequence*, Amer. Math. Monthly **98** (1991), no. 1, 5–20.

[29] Klaus Pinn, *Order and chaos in Hofstadter's Q (n) sequence*, Complexity **4** (1999), no. 3, 41–46.

[30] F. Ruskey, *Fibonacci meets Hofstadter*, Fibonacci Quart. **49** (2011), no. 3, 227–230.

[31] N.J.A. Sloane, *OEIS Foundation Inc.*, The On-Line Encyclopedia of Integer Sequences, 2016. http://oeis.org/.

[32] Stephen M Tanny, *A well-behaved cousin of the Hofstadter sequence*, Discrete Math. **105** (1992), no. 1, 227–239.