

Arithmetische Progressionen oder Ordnung im Chaos

1 Einleitung

Auch noch so chaotische, zufällige Strukturen können gar nicht so unregelmäßig sein, dass es nicht doch hochgradig reguläre Teilstrukturen gibt. Es gibt eine Reihe von Sätzen der Form: Für alle $k \in \mathbb{N}$ gibt es ein $n(k)$, so dass eine beliebige Struktur der Größe $n(k)$ eine reguläre Teilstruktur die Größe k hat.

Für dieses Prinzip nennen wir drei Beispiele. Mit den letzten beiden Beispielen werden wir uns näher befassen. (Einen guten Überblick zu diesem Prinzip gibt [9].)

- Ramseytheorie. Gegeben sei ein vollständiger Graph K_n . Färbt man die Kanten auf beliebige Weise rot oder blau, so gibt es einen vollständigen Teilgraphen K_r , der nur aus einer der beiden Farben besteht, also monochromatisch ist. Man kann zeigen, dass man r von der Größenordnung $r \geq c \log n$ wählen kann. Zur Übung beweise man, dass unter 6 Personen, man immer 3 findet, die sich entweder alle kennen, oder alle nicht kennen. (Wobei „sich kennen“ natürlich als reflexiv und symmetrisch, aber nicht als transitiv angenommen wird).
- B. van der Waerden bewies: Teilt man die natürlichen Zahlen in zwei Klassen (Farben) ein, so muss mindestens eine der Klassen beliebig lange arithmetische Progressionen enthalten. Dies wurde von I. Schur und P. Baudet vermutet.
- E. Szemerédi bewies, dass eine Folge $\mathcal{A} \subset \mathbb{N}$ mit positiver oberer Dichte beliebig lange arithmetische Progressionen haben muss. In anderen Worten, sei $r_k(n)$ die maximale Anzahl einer Menge natürlicher Zahlen in $[1, n]$, die keine Progression der Länge k enthält, so gilt für $k \geq 3$: $r_k(n) = o(n)$. Dies wurde von P. Erdős und P. Turán vermutet.

Das dritte Thema ist mit dem zweiten eng verwandt. Man kann zu Sätzen vom van der Waerden-Typ oft eine sogenannte Dichteverision vom Szemerédi-Typ angeben. Gerade dieser letzte Themenkreis hat in den vergangenen Jahrzehnten eine ungeheure Aufmerksamkeit erhalten. Insbesondere wurden drei Fieldsmedaillen an Leute vergeben, die zu diesem Thema fundamentale Beiträge geleistet haben. (K. Roth, J. Bourgain und T. Gowers).

Für einen Überblick und Verallgemeinerungen vergleiche man auch [28].

In dem Seminar werde ich zunächst die entscheidenden Beweisideen für den Satz von van der Waerden skizzieren und den Satz von Szemerédi für $k = 3$ beweisen.

In einem zweiten Teil fragen wir uns, inwieweit es Umkehrungen gibt. D.h. zu den erhaltenen oberen Schranken der Sätze geben wir komplementäre untere Schranken an.

2 Der Satz von van der Waerden

2.1 Endliche Version

Anstelle des obigen Satzes von van der Waerden betrachten wir folgende Variante:

Satz 1 (van der Waerden, 1928). *Für alle $k, r \in \mathbb{N}$ gibt es eine Zahl $\tilde{W}(k, r)$ mit der Eigenschaft, dass bei einer beliebigen Einteilung der Zahlen $\{1, 2, \dots, \tilde{W}(k, r)\}$ in r Klassen (Farben) mindestens eine der Klassen eine arithmetische Progression der Länge k enthält.*

Wir definieren für festes k und r die kleinstmögliche Zahl $W(k, r) = \min \tilde{W}(k, r)$.

Die ursprüngliche Frage betrachtete nur den Fall $r = 2$, aber beim Beweis stellte es sich heraus, dass alles einfacher wird, wenn man eine stärkere Induktionsbehauptung zugrunde legt. (Lange war kein Beweis bekannt, der den Fall $r = 2$ direkt angeht, vgl. Deuber [10]).

Die Reduktion der unendlichen Mengen mit beliebig langen Progressionen auf endliche Intervalle der Länge $\tilde{W}(k, r)$ mit Progressionen der Länge k folgt mittels „Kompaktheit“ aus der Logik bzw. Mengenlehre. Ein Diagonalenargument zeigt, dass es nicht sein kann, dass die unendliche Version gilt, aber nicht die endliche, vgl. den Aufsatz von van der Waerden [47]. Gilt andersherum die endliche Version, gilt natürlich erst recht die unendliche Version.

2.2 Ein Beispiel

Satz 2. *Es gilt $W(3, 2) = 9$.*

Zum Beweis versuchen wir, eine Progression der Länge 3 zu verhindern. Beweis durch Fallunterscheidung: 1 ist o.B.d.A. rot.

Fall 1.

Sei nun 2 rot, dann ist 3 blau.

r 1 2
b 3

Fall 1.1.

Wir nehmen an, dass 4 rot sei. Dann muss 6 und 7 blau sein, wegen (2,4,6) und (1,4,7). Wegen (5,6,7) muss dann 5 rot sein.

r 1 2 4 5
b 3 6 7

Dann aber kann 8 weder rot sein (2,5,8), noch blau (6,7,8).

Fall 1.2.

Wir nehmen an, 4 sei blau, dann muss wegen (3,4,5) 5 rot sein.

r 1 2 5
b 3 4

Über 6 ist nichts bekannt, man könnte eine weitere Fallunterscheidung machen, aber man will ja auch nicht zu viele Unterfälle betrachten. Wegen (2,5,8) und (1,5,9) müssen 8 und 9 blau sein. Dann aber muss 7 rot sein.

r 1 2 5 7
b 3 4 8 9

Jetzt sieht man, dass 6 weder rot noch blau sein kann.

Andererseits, weil wir die Fallunterscheidung bei 6 nicht vorgenommen haben, sondern gleich benutzt, dass wir Intervalllänge 9 haben, haben wir noch nicht gesehen, ob nicht vielleicht auch Intervalllänge 8 ausgereicht hätte. Das folgende Beispiel zeigt aber $W(3, 2) > 8$.

```
r 1 2          5 6
b          3 4          7 8
```

Wir haben also systematisch Fall 1 behandelt.

Fall 2.

1 ist o.b.d.A. rot, aber 2 ist blau.

```
r 1
b  2
```

Fall 2.1.

Wir nehmen an, 3 sei rot. Es ist dann wegen (1,3,5) 5 blau, und wegen (2,5,8) 8 rot.

```
r 1      3          8
b  2      5
```

Fall 2.1.1.

Wir nehmen an, 4 sei rot. Es ist dann wegen (1,4,7) 7 blau, und wegen (4,6,8) und (5,6,7) kann 6 weder rot noch blau sein.

```
r 1      3 4          8
b  2      5      7
```

Fall 2.1.2.

Wir nehmen an, 4 sei blau. Es ist dann wegen (2,4,6) 6 rot, und wegen (6,7,8) ist 7 blau.

```
r 1      3          6      8
b  2      4 5      7
```

Wegen (3,6,9) und (5,7,9) kann 9 weder rot noch blau sein.

Fall 2.2.

Wir nehmen an, 3 sei blau. Dann muss wegen (2,3,4) 4 rot sein und wegen (1,4,7) 7 blau. Wegen (3,5,7) ist 5 rot.

```
r 1          4 5
b  2 3          7
```

Dann ist aber wegen (4,5,6) 6 blau. Wegen (1,5,9) und (3,6,9) kann 9 weder rot noch blau sein.

```
r 1          4 5          8
b  2 3          6 7
```

ist übrigens ein zweites Beispiel für $W(3, 2) > 8$.

2.3 Der allgemeine Fall

Es ist folgendes bekannt:

$$W(2, r) = r+1, W(3, 2) = 9, W(3, 3) = 27, W(3, 4) = 76, W(4, 2) = 35, W(5, 2) = 178, W(6, 2) > 230.$$

Die Funktion $W(k, r)$ wächst so stark und nicht vorhersehbar an, dass man dies nicht mehr allgemein analog durch Fallunterscheidungen bearbeiten kann. Die besten bekannten oberen Schranken sind riesengroß. Das Problem ist, dass man eine allgemeine Methode benötigt. Man verwendet doppelte Induktion und dabei werden die vorher berechneten Werte rekursiv aufgerufen.

2.3.1 $W(3, 2) \leq 325$

Wir zeigen mit einer anderen Methode, (die verallgemeinerungsfähig ist), $W(3, 2) \leq 325$. Erneut versuchen wir, eine Progression der Länge 3 zu verhindern. Wir unterteilen $[1, 325]$ in 65 Blöcke B_i der Länge 5:

$$[1, 5] \cup [6, 10] \cup \dots \cup [321, 325].$$

Vgl. Bild 1.

Es gibt $2^5 = 32$ Möglichkeiten, die 5er Blöcke mit 2 Farben zu färben. Unter den ersten 33 B_i sind also zwei mit gleicher Färbung, nennen wir diese B_a und B_b . Wir beachten, dass auch B_{2b-a} einer der 65 Blöcke ist.

$$B_a = \{5(a-1) + 1, 5(a-1) + 2, 5(a-1) + 3, 5(a-1) + 4, 5(a-1) + 5\}.$$

Unter den ersten drei Zahlen gibt es zwei, die die gleiche Farbe (o.B.d.A. rot) haben, sagen wir j und $j + d$, wobei $d \in \{1, 2\}$. Falls $j + 2d$ auch rot wäre, sind wir fertig, $j + 2d$ ist also blau. Wir haben übrigens die Blöcklänge 5 gewählt, damit $j + 2d$ im gleichen Block liegt.

Vgl. Bild 2.

Von den 5 Elementen interessieren uns also nur drei in Progression, $j, j + d, j + 2d$. B_a sieht also vereinfacht wir rrb aus. Genauso sind in B_b die drei Zahlen $5(b-a) + \{j, j + d, j + 2d\}$ als rrb gefärbt. Betrachten wir nun B_{2b-a} , insbesondere die Zahl $10(b-a) + j + 2d$. Diese kann wegen $(j, 5(b-a) + j + d, 10(b-a) + j + 2d)$ nicht rot sein, und wegen $(j + 2d, 5(b-a) + j + 2d, 10(b-a) + j + 2d)$ nicht blau sein. Diese Farbfokussierung auf eine Zahl zeigt, dass $[1, 325]$ eine monochromatische Progression der Länge drei enthalten muss.

Vgl. Bild 3.

2.3.2 $W(3, 3) < \infty$

Satz 3.

$$W(3, 3) \leq 7(2 \cdot 3^7 + 1)(2 \cdot 3^{7(2 \cdot 3^7 + 1)} + 1).$$

Diese Zahl ist riesengroß, ungefähr $10^{14.600}$, wo man doch $W(3, 3) = 27$ nachweisen kann. Immerhin funktioniert dies mit einer allgemeinen Methode, und man sieht in diesem Abschnitt den allgemeinen Beweis durchschimmern.

Wir teilen das Intervall $[1, 7(2 \cdot 3^7 + 1)(2 \cdot 3^{7(2 \cdot 3^7 + 1)} + 1)]$ in $(2 \cdot 3^{7(2 \cdot 3^7 + 1)} + 1)$ Blöcke B_i der Länge $l = (7 \cdot 3^7 + 1)$ ein. Es gibt 3^l viele Möglichkeiten, diese Blöcke mit drei Farben zu färben. Es sei $z = 7(2 \cdot 3^7 + 1)$. Unter den ersten $3^z + 1$ Blöcken B_i müssen also zwei identisch gefärbt sein, B_{i_1} und $B_{i_1+d_i}$. Man beachte, dass $B_{i_1+2d_i}$ definiert ist, daher also die Anzahl der Blöcke. Beim letzten Schritt benutzen wir, dass $W(2, r)$ für riesengroßes r bekannt ist. Allgemein, beim Induktionsschritt schließt man von $W(k-1, r')$ für große r' und $W(k, r-1)$ auf $W(k, r)$.

Wir zerlegen die Blöcke B_i in $2 \cdot 3^7 + 1$ Teilblöcke $B_{i,j}$ der Länge 7. Es gibt 3^7 Färbungen dieser 7er Blöcke, also gibt es unter den ersten $3^7 + 1$ Blöcken zwei Blöcke B_{i_1, i_2} und $B_{i_1, i_2 + d_2}$ die gleich gefärbt sind. Unter den ersten 4 Zahlen des ersten 7er Blocks B_{i_1, i_2} kommt eine Farbe doppelt vor. O.B.d.A. sind also i_3 und $i_3 + d_3$ beide rot. Hierbei ist $d_3 \in \{1, 2, 3\}$. Beachte $\{i_3, i_3 + d_3, i_3 + 2d_3\} \in B_{i_1, i_2}$. Es ist also o.B.d.A.: rrb. Da

$$\{i_3 + 7d_2, i_3 + 7d_2 + d_3, i_3 + 7d_2 + 2d_3\} \in B_{i_1, i_2 + d_2}$$

identisch gefärbt ist, liegt auch hier der Fall rrb vor. Analog wie im vorigen Abschnitt über $W(3, 2)$ folgt, dass $i_3 + 14d_2 + 2d_3$ weder rot noch blau sein kann, also gelb sein muss. Soweit haben wir nur den großen Block B_{i_1} behandelt. In Kurzform also rrb rrb xxg. Der Block B_{i_2} ist identisch gefärbt. Die konkreten Zahlen sind (Hinweis, ab jetzt hat (zumindest die erste Auflage) [26] Ungenauigkeiten):

$$\{i_3 + ld_1, i_3 + d_3 + ld_1, i_3 + 2d_3 + ld_1\} \in B_{i_1 + d_1, i_2}.$$

$$\{i_3 + ld_1 + 7d_2, i_3 + d_3 + ld_1 + 7d_2, i_3 + 2d_3 + ld_1 + 7d_2\} \in B_{i_1 + d_1, i_2 + d_2}.$$

Und analog

$$i_3 + 2d_3 + ld_1 + 14d_2 \in B_{i_1 + d_1, i_2 + d_2}.$$

Das ergibt also schematisch (vergleiche auch Bild 4):

$$\begin{array}{ccccccc} r & r & b & r & r & b & x & x & g \\ r & r & b & r & r & b & x & x & g \\ x & x & x & x & x & x & x & x & ? \end{array}$$

Wenn wir nun auf

$$m = i_3 + 2d_3 + 2ld_1 + 14d_2$$

fokussieren, so kann m wegen

$$i_3, i_3 + d_3 + ld_1 + 7d_2, m = i_3 + 2d_3 + 2ld_1 + 14d_2$$

nicht rot sein, wegen

$$i_3 + 2d_3, i_3 + 2d_3 + ld_1 + 7d_2, m = i_3 + 2d_3 + 2ld_1 + 14d_2$$

nicht blau sein, und wegen

$$i_3 + 2d_3 + 14d_2, i_3 + 2d_3 + ld_1 + 14d_2, m = i_3 + 2d_3 + 2ld_1 + 14d_2$$

nicht gelb sein.

(Folgendes ist etwas unpräzise: Durch das rekursive Zurückgreifen auf Färbungen ganzer Blöcke, erhält man riesige Zahlen. Die Blockreihe hat zwar nur Länge $k-1$, dafür benötigt man aber für die Schranke für $W(k, r)$ die Schranken für $W(k-1, r')$ mit sehr großem r' . Die obigen Schranken $W(k, 2)$ wachsen bereits schneller als die Ackermannfunktion, die wiederum ein Beispiel einer nicht primitiv rekursiven Funktion ist. Diese ist wie folgt definiert. Es sei

$$\begin{aligned} A(0, n) &= n + 1, \\ A(m + 1, 0) &= A(m, 1), \\ A(m + 1, n + 1) &= A(m, A(m + 1, n)) \end{aligned}$$

Daraus folgt dann z.B.:

$$A(0, n) = n + 1,$$

$$A(1, n) = 2 + (n + 3) - 3,$$

$$A(2, n) = 2(n + 3) - 3,$$

$$A(4, n) = 2^{2^{\dots^2}} - 3 \text{ (Turm der Höhe } n + 3 \text{ etc. Die Ackermannfunktion ist die Diagonale } A(n, n).$$

Shelah (siehe [35], aber auch die zweite Auflage von [26]) hat gezeigt, dass man nicht die Ackermannfunktion, sondern nur die 5.te oder 6.te Stufe benötigt. Aber auch diese Schranken sind ungeheuer groß.

In diesem Jahr nun erschien die Arbeit von Gowers über die Verbesserung des Satzes von Szemerédi. Während Szemerédi in seinem Beweis van der Waerdens Satz (mit den damals bekannten Werten) benutzte, und daher keine neuen Schranken an $W(k, r)$ erhielt, dafür aber schlechte Schranken für $r_k(n)$, vermeidet Gowers den Satz von van der Waerden. Aus guten Schranken für $r_k(n)$ kann er dann auch gute Schranken für $W(k, r)$ erhalten. Gowers erhielt für seine Schranken einen von Ron Graham ausgelobten Preis von \$1000. Grahams Vermutung besagte, dass für $W(k, 2)$ ein Turm von 2er-Potenzen der Höhe k ausreicht. Gowers bewies also für große k etwas viel besseres, aber für $5 \leq k \leq 8$ ist die Vermutung meines Wissens noch offen.)

Satz 4 (Gowers, 2001).

$$W(k, 2) \leq 2^{2^{2^{2^{k+9}}}}.$$

2.4 Untere Schranken

Die komplementären unteren Schranken sind sehr weit von den oberen Schranken entfernt.

Satz 5 (Berlekamp, 1968). Für primes p gilt $W(p + 1, 2) \geq p 2^p$.

(Argumentation in endlichen Körpern. Der Beweis geht so ähnlich, wie in der Codierungstheorie, wo man bestimmte Codes konstruiert.)

Wir beweisen etwas weniger. Wir verwenden ein probabilistisches Argument unter Benutzung des Lovász Local Lemma.

Satz 6.

$$W(k, 2) \geq \frac{2^k}{2ek} (1 + o(1)).$$

Seien A_i, B_i Ereignisse eines Wahrscheinlichkeitsraums Ω . Es sei A unabhängig von B_1, \dots, B_r , in folgendem stärkeren Sinne von Unabhängigkeit: Für alle nichtleeren Teilmengen $J \subseteq [1, r]$ gilt

$$p\left(A \wedge \left(\bigwedge_{j \in J} B_j\right)\right) = p(A)p\left(\bigwedge_{j \in J} B_j\right).$$

Wir nennen dann A von den B_i wechselseitig unabhängig (mutually independent).

Es bezeichne \mathcal{S} die Menge aller arithmetischen Progressionen $S \subset [1, n]$ der Länge k . Man beachte, dass (für festes k) es etwa $c_k n^2$ viele Progressionen gibt, da die Progressionen der Länge k , also $\{a + ib : 0 \leq i \leq k - 1\}$, ja durch zwei Parameter a und b beschrieben werden können.

$$c'_k n^2 \leq |\mathcal{S}| \leq n^2.$$

Wir definieren auf \mathcal{S} den Abhängigkeitsgraphen G . Die Ecken des Graphen sind die Mengen $S \in \mathcal{S}$, also $E = \mathcal{S}$. Hierbei sind $S \in E$ und $T \in E$ durch eine Kante verbunden, wenn $S \cap T \neq \emptyset$. (Achtung: im Vortrag hatte ich aus Versehen $S \cap T = \emptyset$ gesagt!) Der Graph verbindet also arithmetische Progressionen, die sich schneiden. Da dies für großes n und kleines k recht selten vorkommt, ist der Graph also recht leer. Eine Progression, die eine zweite Progression nicht schneidet, ist von der zweiten Progression wechselseitig unabhängig.

Um das „wechselseitig“ brauchen wir uns nicht kümmern, es ist halt nur Voraussetzung des benötigten Hilfsmittels.

Wir färben nun das Intervall $[1, n]$ mit zwei Farben zufällig. Die Zahl i ist mit Wahrscheinlichkeit $p = \frac{1}{2}$ rot bzw. blau.

Wir bezeichnen mit A_S das Ereignis, dass die arithmetische Progression S monochrom ist und mit \bar{A}_S das komplementäre Ereignis. Es gilt also $p(A_S) = \frac{2}{2^k}$ und $p(\bar{A}_S) = 1 - \frac{1}{2^{k-1}}$.

Wir fragen uns, für welches $n(k)$ wir

$$p\left(\bigwedge_{S \in \mathcal{S}} \bar{A}_S\right) > 0$$

nachweisen können. Denn dann gibt es eine Färbung des Intervalls, in der **keine** monochromatische arithmetische Progression existiert. Somit würde dieses n eine untere Schranke an $W(k, 2)$ liefern.

Wir zitieren nach Alon und Spencer [2]:

Lemma 1 (Lovász Local Lemma). *Seien A_1, A_2, \dots, A_r Ereignisse eines Wahrscheinlichkeitsraumes. Ein gerichteter Graph $D(V, E)$ auf der Menge der Kanten $V = \{1, \dots, r\}$ wird Abhängigkeitsgraph der Ereignisse A_1, \dots, A_r genannt, wenn für alle i , ($1 \leq i \leq r$), das Ereignis A_i wechselseitig unabhängig von allen Ereignissen $\{A_j : (i, j) \notin E\}$ ist. Sei also $D(V, E)$ ein derartiger gerichteter Graph zu den genannten Ereignissen. Wir nehmen an, dass es reelle Zahlen $0 \leq x_i < 1$ gibt, und dass für alle $1 \leq i \leq r$*

$$p(A_i) \leq x_i \prod_{(i,j) \in E} (1 - x_j)$$

gilt. Dann gilt

$$p\left(\bigwedge_{i=1}^r \bar{A}_i\right) \geq \prod_{i=1}^r (1 - x_i).$$

(Hinweis: Mir scheint, dass bei dieser Definition Ereignisse, die durch Kanten verbunden sind, dennoch unabhängig sein können. Der Graph ist also zunächst nicht eindeutig definiert. Für Anwendungen ist es aber günstig, ihn so klein wie möglich zu wählen, und nicht willkürlich Kanten stehen zu lassen.) Man beachte, dass die Voraussetzung an die Wahrscheinlichkeiten lokal sind, dass die Aussage aber global ist. Gerade deswegen ist das Lemma nützlich. Für viele Anwendungen reicht es bereits, dass $p(\bigwedge_{i=1}^r \bar{A}_i) > 0$ gilt.

Das Lemma ist recht allgemein, oft braucht man nur einen symmetrischen Spezialfall. Dieser ist besonders nützlich, wenn der Graph nicht sehr dicht ist, was ja in unserer Situation der Fall ist.

Korollar 1. *Es seien A_1, \dots, A_r Ereignisse wie oben. Die A_i seien von fast allen anderen A_j , mit höchstens d Ausnahmen, wechselseitig unabhängig. Der Abhängigkeitsgraph hat also maximalen Grad d (maximale Anzahl von Kanten, die von einer Ecke ausgeht). Es sei $p(A_i) \leq p$ für $1 \leq i \leq r$. Falls*

$$ep(d+1) \leq 1$$

gilt, dann gilt auch

$$p\left(\bigwedge_{i=1}^r \bar{A}_i\right) > 0.$$

Zum Beweis des Korollars: Für $d = 0$ (leerer Graph) ist alles klar, für $d \geq 1$ setze man für $x_i = \frac{1}{d+1} < 1$. Man beachte $\left(1 - \frac{1}{d+1}\right)^d > \frac{1}{e}$.

Bei der Anwendung auf die van der Waerden Zahlen sehen wir nun: Eine arithmetische Progression der Länge k schneidet höchstens nk andere Progressionen. (Gehen wir zu dem Schnittpunkt, von dort kann es höchstens n andere Progressionen geben. Andererseits kommen bei jeder Progression k Punkte als Schnittpunkte in Frage.)

Um das Korollar anwenden zu können, muss also $p = \frac{1}{2^{k-1}} < \frac{1}{e^{(nk+1)}}$ gelten. Dies ist für $n < \frac{2^k}{2ek}(1 + o(1))$ gegeben.

3 Der Satz von Roth bzw. der Satz von Szemerédi für $k = 3$

Erdős und Turán [14] betrachteten für sehr kleine Intervalle $[1, n]$ die maximale Anzahl $r_3(n)$ von Zahlen, die keine arithmetische Progression der Länge 3 enthalten. So bewiesen sie für $n \geq 8$, dass $r_3(2n) \leq n$ gilt, und verbesserten dies schrittweise für genügend große n zu $r_3(n) < \left(\frac{3}{8} + \varepsilon\right)n$. Sie vermuteten, dass man so im Prinzip zu immer besseren Schranken komme, dass also $r_3(n) = o(n)$ gilt.

Insbesondere hoffte man anfangs, durch gute Schranken bei diesem Problem auch die Schranken im van der Waerden Problem zu verbessern.

Eine weitere Motivation war, dass man auch die Existenz beliebig langer Progressionen in der Menge der Primzahlen nachweisen wollte. Falls man nämlich nachweisen kann, dass für die maximale Anzahl $r_k(n)$ von Zahlen im Intervall $[1, n]$, die keine arithmetische Progression der Länge k enthalten, $r_k(n) = o\left(\frac{n \log \log n}{\log n}\right)$ gilt, so folgt, dass es Primzahlprogressionen beliebiger Länge geben muss. (Für $r_k(n) = o\left(\frac{n}{\log n}\right)$ folgt dies sofort aus den Tschebyscheffschranken $\pi(n) \gg \frac{n}{\log n}$. Für die Verfeinerung vergleiche man [17].)

Erdős setzte einen Preis von \$1000 für folgende Vermutung aus:

Vermutung 1.

$$r_k(n) = o(n).$$

Auch wenn die Vermutung nach Erdős und Turán benannt ist, so bekommt man beim Lesen von [14] den Eindruck, dass die Vermutung vielleicht eher von Szekeres stammt.

Darüber hinaus vermutete er:

Vermutung 2 (Erdős, \$3000). Sei (a_i) eine monoton wachsende Folge und $\sum_i \frac{1}{a_i}$ sei divergent. Dann enthält die Folge arithmetische Progressionen beliebiger Länge.

Diese Vermutung impliziert natürlich den Fall der Primzahlen.

Die Erdős-Turán Vermutung wurde von Roth für $k = 3$ bewiesen. Der Beweis benutzte die Hardy-Littlewoodsche Kreismethode.

Satz 7 (Roth, [32]). Es sei \mathcal{A} eine Menge positiver Zahlen mit positiver oberer Dichte. Dann enthält \mathcal{A} eine arithmetische Progression der Länge 3.

Sei andersherum $r_3(n)$ die größte Anzahl von Zahlen im Intervall $[1, n]$, die keine arithmetische Progression der Länge 3 haben, dann gilt $r_3(n) = o(n)$. Genauer gilt sogar $r_3(n) = O\left(\frac{n}{\log \log n}\right)$.

Dies wurde wie folgt verbessert.

Satz 8 (Szemerédi [42], Heath-Brown [27]). *Es gibt eine positive Konstante c , so dass $r_3(n) = O\left(\frac{n}{(\log n)^c}\right)$ gilt.*

Das derzeit stärkste Ergebnis ist:

Satz 9 (Bourgain, [8]). *Es gilt $r_3(n) = O\left(\frac{n}{\left(\frac{\log n}{\log \log n}\right)^{1/2}}\right)$.*

Folgende untere Schranken sind bekannt. Man beachte, dass zwischen der oberen und unteren Schranke noch viel Raum für Verbesserungen ist.

Satz 10 (Behrend, [4]). *Es gibt eine positive Konstante c , so dass gilt:*

$$r_3(n) \geq \frac{n}{\exp(c(\log n)^{1/2})}.$$

(Der Beweis ist geometrisch sehr anschaulich. Man legt eine Menge von Punkten auf eine (hochdimensionale) Sphäre. Geraden durch diese Punkte schneiden die Menge natürlich in keinem dritten Punkt. Auf Details gehen wir hier nicht ein.)

Man hätte vermuten können, dass es für Mengen mit mindestens N^α vielen Elementen zumindest für $\alpha > \frac{1}{2}$ immer Progressionen der Länge 3 gibt. Es gibt dann ja $N^{2\alpha}$ viele Paare, und die Chance, dass bei derart vielen Paaren wenigstens einmal das dritte Element dann auch zur Menge gehört, ist gut. Die untere Schranke zeigt also, dass der probabilistisch typische Fall und der extremale Fall weit voneinander entfernt sind. Ob der extremale Fall sich aber durch eine geometrische Konstruktion beschreiben läßt, ist natürlich keineswegs klar.

(Sollte jemand also einen probabilistischen Beweis von 200 Seiten Länge für eine besser untere Schranke angeben, wird der Beweis wohl schon alleine deswegen falsch sein!)

Soweit zum Fall $k = 3$. Szemerédi bewies den Fall $k = 4$ mit kombinatorischen Methoden. Roth gab später einen analytischen Beweis an. Diese Beweise sind leider sehr kompliziert. Schließlich bewies Szemerédi den allgemeinen Fall und verdiente sich damit den ausgesetzten Preis. (Meines Wissens ist \$1000 der höchste Erdős-Preis der jemals ausgezahlt wurde. (Die Wirtingschen \$10¹⁰ wurden ja natürlich nicht ausgezahlt.)

Satz 11 (Szemerédi, [41]). *Es gilt $r_k(n) = o(n)$.*

Da Szemerédi die schlechten Schranken von dem van der Waerden Beweis verwendet, erhält er keine obere Schranke vom Typ $r_k(n) = O\left(\frac{n}{\log \log \dots \log n}\right)$. Seine Schranke ist nur so extrem wenig besser als $o(n)$, dass dies niemand genau ausgearbeitet hat.

Später gab Fürstenberg einen ergodentheoretischen Beweis an. Für Leute, die sich mit den Methoden auskennen, ist dieser Beweis sicher leichter. Es gab auch einige Verllgemeinerungen von Szemerédi's Satz, z.B. auf mehrdimensionale Strukturen oder auf polynomiale Werte, die meistens zunächst mittels der ergodentheorie bewiesen wurden. Leider liefert diese Methode abger grundsätzlich nur ein $o(n)$ und keine besseren Schranken. Die Arbeit von Gowers, die im wesentlichen Methoden der harmonischen Analyse verwendet, liefert für alle k

$$r_k(n) = O_k\left(\frac{n}{(\log \log n)^{c_k}}\right) \quad !!$$

Dies war ja nicht einmal für $k = 4$ bekannt! Gowers beschreibt Roths Arbeit so: es gelingt Unterprogressionen der Länge \sqrt{N} zu finden, auf denen die Dichte der Menge ein bisschen zugenommen hat. Wenn man das ein paar Mal iteriert, nimmt die Dichte weiter zu. Allerdings

gibt die starke Abnahme der Länge der Folge eine Grenze an, wie oft man iterieren kann. Diese Idee verwendet Gowers, allerdings sehr viel effizienter.

Wie dem auch sei, auch Szemerédis Beweis wird überleben. Er enthält nämlich zwei wichtige Lemmata, die wiederum selber zum Forschungsgegenstand und wichtigen Hilfsmittel bei anderen Fragen wurden: das Regularity Lemma der Graphentheorie (siehe [7]) und das Cube Lemma. Gowers [19] bewies übrigens auch, und das war unerwartet, dass die schwachen Abschätzungen beim Regularity Lemma nicht wesentlich verbessert werden können. Er bewies also, dass es untere Schranken vom Turmexponentialtyp gibt. Dadurch war klar, dass es nicht reicht, Szemerédis Beweis zu verfeinern, um zu einer wesentlich besseren Schranke zu kommen.

Wir begnügen uns beim Beweis mit $r_3(n) = o(n)$. Den Beweis führen wir nach [26].

Wir nennen $M \in \mathcal{A}$ einen Würfel, wenn es einen Aufpunkt a und Seitenlängen d_i derart gibt, dass gilt:

$$M(a; d_1, \dots, d_l) = \left\{ a + \sum_{i=1}^l \varepsilon_i d_i : \varepsilon_i \in \{0, 1\} \right\}.$$

(Man beachte, dass $\mathcal{A} \subset \mathbb{N}$ gilt, dass der Würfel also nur eine gedachte mehrdimensionale Struktur auf der eindimensionalen Zahlengeraden ist.)

Lemma 2 (Cube Lemma). *Sei $\mathcal{A} \subset [1, n]$ mit $|\mathcal{A}| \geq cn$. Dann gibt es einen Würfel der Dimension $l = \log \log n + O(1)$ mit $M \subset \mathcal{A}$.*

Beweisskizze:

Es sei $|\mathcal{A}| = \alpha_0$. Es gibt $\binom{\alpha_0}{2}$ Differenzen $a_1 - a_2, a_1 \in \mathcal{A}, a_2 \in \mathcal{A} \setminus \{a_1\}$, die positiv sind. Da wegen $\mathcal{A} \subset [1, n]$ nur $n - 1$ verschiedenen Differenzen in Frage kommen, muss es eine Differenz d_1 geben, die mindestens $\frac{\binom{\alpha_0}{2}}{n - 1}$ mal vorkommt. Wir bilden

$$\mathcal{A}_1 = \{a \in \mathcal{A} : a + d_1 \in \mathcal{A}\}.$$

Es gilt $|\mathcal{A}_1| \geq \alpha_1 := \frac{\binom{\alpha_0}{2}}{n - 1}$. Analog finden wir eine häufige Differenz d_2 , so dass

$$\mathcal{A}_2 = \{a \in \mathcal{A}_1 : a + d_2 \in \mathcal{A}_1\} = \{a + \varepsilon_1 d_1 + \varepsilon_2 d_2 \in \mathcal{A}, \varepsilon_i \in \{0, 1\}\}.$$

Auf diese Weise schachtelt man immer größere Würfel ineinander:

$$\mathcal{A}_l \subset \mathcal{A}_{l-1} \subset \dots \subset \mathcal{A}_1 \subset \mathcal{A}$$

mit $d_i + \mathcal{A}_i \subset \mathcal{A}_{i-1}$, $|\mathcal{A}_i| \geq \alpha_i$, wobei $\alpha_i = \frac{\binom{\alpha_{i-1}}{2}}{n - 1}$ gilt.

Es ist $\alpha_1 \geq \frac{cn(cn - 1)}{2(n - 1)} \approx \frac{c^2 n}{2}$ und analog $\alpha_i \approx \frac{c^{2^i} n}{2^{2^i - 1}}$. (Genaugenommen verliert man in jedem Schritt noch ein ganz kleines bisschen. Es ist etwas aufwendig, dies ganz genau zu machen. Man könnte dies mit ε ausdrücken.) Falls $\alpha_i \geq 1$ kann man auch im i -ten Schritt noch ein d_i wählen. Die Schrittzahl ist daher wegen $\log \log \left(\frac{2}{c}\right)^{2^l} \approx \log \log n$ ungefähr $l \approx \frac{\log \log n}{\log 2} + \log \log \frac{2}{c}$, also für großes n auf jeden Fall größer als $\log \log n$.

Es ist erstaunlich einfach nachzuweisen, dass der Grenzwert $\lim_{n \rightarrow \infty} \frac{r_k(n)}{n}$ existiert. Das Problem ist nachzuweisen, dass er gleich Null ist.

Lemma 3 (Subadditivität). *Der Grenzwert $\lim_{n \rightarrow \infty} \frac{r_k(n)}{n}$ existiert. Darüber hinaus gilt $\lim_{m \rightarrow \infty} \frac{r_k(m)}{m} \leq \frac{r_k(n)}{n}$.*

Proof. Aus der Definition von $r_k(n)$ folgt unmittelbar, dass

$$r_k(m+n) \leq r_k(m) + r_k(n),$$

denn bei dem längeren Intervall können ja noch Progressionen auftreten, die in keinem der Teile auftreten. Es folgt also $r_k(2n) \leq 2r_k(n)$ bzw. induktiv das Analogon für andere Vorfaktoren. Wir zerlegen m wie folgt: $m = \left[\frac{m}{n}\right]n + (m - \left[\frac{m}{n}\right]n)$. Man beachte, dass $(m - \left[\frac{m}{n}\right]n) < n$ gilt. Daher ist

$$r_k(m) \leq r_k(n) \left[\frac{m}{n}\right] + r_k(m - n \left[\frac{m}{n}\right]) \leq r_k(n) \frac{m}{n} + n.$$

Daraus folgt

$$\frac{r_k(m)}{m} \leq \frac{r_k(n)}{n} + \frac{n}{m}.$$

Dies gilt für alle m . Also gilt auch

$$\limsup_{m \rightarrow \infty} \frac{r_k(m)}{m} \leq \frac{r_k(n)}{n}.$$

Da dies nun auch für alle n gilt, folgt auch

$$\limsup_{m \rightarrow \infty} \frac{r_k(m)}{m} \leq \liminf_{n \rightarrow \infty} \frac{r_k(n)}{n}.$$

Der Grenzwert existiert also, d.h. es ist $\limsup = \liminf$. Dann aber folgt wegen

$$\lim_{m \rightarrow \infty} \frac{r_k(m)}{m} \leq \frac{r_k(n)}{n}$$

auch die zweite Behauptung. □

Proof. Wir beweisen nun den Satz von Roth. Angenommen es sei $c = \lim_{n \rightarrow \infty} \frac{r_3(n)}{n} > 0$, so gilt wegen der Subadditivität $r_3(n) \geq cn$. Wir setzen $\varepsilon = 10^{-10}c^2$, und n_0 sei groß. Für alle genügend großen $n \geq n_1 \geq n_0$ gilt

$$c \leq \frac{r_3(n)}{n} \leq c + \varepsilon.$$

Hinweis: später benötigen wir, dass $\frac{c^2 \log \log n}{100} > n_0$ gilt. Es sei nun $\mathcal{A} \subset [1, n]$ eine Menge mit $|\mathcal{A}| = r_3(n) \geq cn$, aber ohne Progression der Länge 3. Die Folge muss einigermaßen gleichmäßig verteilt sein. Wir überlegen uns kurz, dass \mathcal{A} im Intervall $(0.49n, 0.5n)$ mindestens Dichte $\frac{\varepsilon}{2}$ haben muss, (genaugenommen sogar fast c). Das Intervall $[1, 0.49n]$ hat höchstens $(c + \varepsilon)0.49n$ Elemente mit \mathcal{A} gemeinsam. Das Intervall $[0.5n, n]$ hat höchstens $(c + \varepsilon)n - \frac{\varepsilon}{2}n$ Elemente mit \mathcal{A} gemeinsam. (An der oberen Grenze n eine obere Schranke verwenden und davon eine untere Schranke für die untere Grenze $0.5n$ abziehen.) Da es insgesamt in $[1, n]$ aber mindestens cn viele Elemente in \mathcal{A} gibt, muss also $(0.49n, 0.5n) \cap \mathcal{A}$ mindestens

$$cn - (c + \varepsilon)0.49n - \left((c + \varepsilon)n - \frac{c}{2}n\right) = \frac{cn}{100} - \frac{149\varepsilon n}{100} \geq \frac{cn}{2 \times 100}$$

viele Elemente enthalten.

Dieses war die erste Aufspaltung des Intervalls in Teile. Es werden noch einige folgen. Wir spalten $(0.49n, 0.5n)$ in Intervalle der Größe $[\sqrt{n}]$ auf. Mindestens eines dieser Intervalle muss Dichte $\frac{c}{2}$ haben.

Das Cube Lemma garantiert, dass es einen Würfel gibt mit folgenden Eigenschaften:

1. $M(a; d_1, \dots, d_l) \subset \mathcal{A} \cap (0.49n, 0.5n)$. (Genauer: sogar in diesem Teilintervall der Länge $[\sqrt{n}]$.)
2. $l = \log \log[\sqrt{n}] + O(1) = \log \log n + O(1)$.
3. $d_i \leq 2\sqrt{n}$ für $i = 1, \dots, l$.

Wir setzen $M_{-1} = \{a\}$, wobei ja $a \in (0.49n, 0.5n)$. Schrittweise bauen wir wie im Beweis des Cube Lemma die Kette von Würfeln auf:

$$M_i = M(a; d_1, \dots, d_i), 0 \leq i \leq l.$$

Da in \mathcal{A} keine arithmetische Progression der Länge 3 vorkommt, ist mit $x, m \in \mathcal{A}$ automatisch $y = 2m - x \notin \mathcal{A}$. Da nun \mathcal{A} einen großen Würfel enthält, kann man daraus eine große Struktur von derart verbotenen Elementen y ableiten. Wir notieren in jedem Schritt die verbotenen Elemente.

$$N_i = \{2m - x : x \in (\mathcal{A} \cap [1, a]), m \in M_i\}.$$

Da ja in \mathcal{A} gerade diese Elemente verboten sind, gilt $\mathcal{A} \cap N_i = \emptyset$. Es gilt also

$$|N_i| \geq |N_{-1}| = |\{2a - x : x \in \mathcal{A} \cap [1, a]\}| = |\mathcal{A} \cap [1, a]| \geq 0.49cn.$$

Da $M_{i+1} = M_i \cup (M_i + d_i)$ gilt analog $N_{i+1} = N_i \cup (N_i + 2d_i)$. Da die $|N_i|$ monoton wachsen, aber natürlich nach oben beschränkt sind, $|N_i| \leq n$, muss es einen Schritt $i \rightarrow i + 1$ geben, wo

$$|N_{i+1} \setminus N_i| < \frac{n}{l}.$$

(Natürlich gilt sogar etwas Strengeres, nach obigen Bemerkungen über N_{-1} .)

Eine arithmetische Progression $\mathcal{B} = \{x, x + 2d_i, x + 4d_i, \dots, x + s(2d_i)\}$ mit Differenz nennen wir einen Block. Ist \mathcal{B} in N_i aber, $x + (s + 1)2d_i \in N_{i+1} \setminus N_i$, so nennen wir den Block einen maximalen Block. Man beachte, dass also maximale Blöcke in N_i und Elemente in $N_{i+1} \setminus N_i$ bijektiv zugeordnet sind. Wir können also Blöcke zählen, indem wir Elemente zählen. Außerdem ist die Anzahl der Blöcke im Komplement $[1, n] \setminus N_i$ gleich (bis auf evtl. einen am Rand). Da also $|N_{i+1} \setminus N_i| < \frac{n}{l}$ gilt, wissen wir jetzt, dass die Anzahl der maximalen Blöcke in $[1, n] \setminus N_i$ ebenfalls höchstens $\frac{n}{l} + 1$ beträgt.

Nun, kommt eine letzte Aufteilung. Wir teilen $[1, n]$ in $2d_i$ Klassen modulo $2d_i$ ein. Dann haben wir $[1, n] \setminus N_i$ in höchstens $\frac{n}{2d_i} + 2d_i \sim \frac{n}{2d_i}$ Blöcke geteilt, (da $d_i \leq 2\sqrt{n}$ und $l \approx \log \log n$). Wir zählen zunächst die Anzahl aller Elemente in kurzen Blöcken mit Länge von höchstens $\frac{c^2 l}{100}$. Die Anzahl beträgt also insgesamt höchstens $\frac{c^2 n}{100} (1 + o(1))$. Die anderen, langen Blöcke sind lang genug, um zu garantieren, dass die Dichte der Menge \mathcal{A} auf ihnen höchstens $c + \varepsilon$ beträgt. (Denn wenn ein langer Block zu große Dichte hat, kann man ihn (und entsprechend \mathcal{A} ja durch eine affine Transformation auf ein Intervall $[1, r]$ abbilden. Wenn r groß genug ist, hat dort \mathcal{A}' höchstens Dichte $c + \varepsilon$. Ganz oben wählten wir $n \geq n_1 \geq n_0$ groß genug, damit dieses Argument keine Probleme macht.)

Wir zählen nun abschließend alle Elemente in \mathcal{A} und erhalten einen Widerspruch:

$$\begin{aligned}
|\mathcal{A}| &= |\mathcal{A} \cap [1, n]| \\
&= |\mathcal{A} \cap ([1, n] \setminus N_i)| \quad \text{da } \mathcal{A} \cap N_i = \emptyset \\
&= \text{Anzahl der Elemente in kurzen und langen Blöcken} \\
&< \frac{c^2 n}{100}(1 + o(1)) + (c + \varepsilon)(n - |N_i|) \\
&= \frac{c^2 n}{100}(1 + o(1)) + cn - 0.49c^2 n + \varepsilon n - \varepsilon |N_i| + o(n) \\
&< cn
\end{aligned}$$

da ε oben sehr klein gewählt war; $\varepsilon = \frac{c^2}{4}$ sollte auch bereits ausreichen. Dieser Widerspruch beweist also, dass $r_3(n) = o(n)$ gilt. \square

Wir hoffen, dass dem Leser durch die parallele Betrachtung von dem Satz von van der Waerden und dem Satz von Roth bzw. Szemerédi auch die gegenseitige Abhängigkeit aufgefallen ist. Der Satz von Szemerédi ist der stärkere Satz. Man bezeichnet diese Form von Satz auch als **Dichteverision** der endlichen van der Waerden Version. Aus dem Satz von Szemerédi folgt der Satz von van der Waerden. Beweistechnisch benutzte Szemerédi allerdings den Satz von van der Waerden.

Es gibt in der kombinatorischen Zahlentheorie eine ganze Reihe von Arbeiten, die andere Muster als Progressionen vermeiden. Man kann ja eine Progression der Länge 3 auch als Lösung der Gleichung $a_1 + a_3 = 2a_2$ mit $a_1, a_2, a_3 \in \mathcal{A}$ auffassen. Derartige Gleichungen kann man natürlich weit verallgemeinern, wobei dann die Progressionsstruktur verloren geht. Hierzu gibt es Arbeiten von Rado (die Hunderte weiterer Arbeiten hervorrief, Chung, Graham, Ruzsa, Schoen, Łuczak und vielen anderen. Bei der umfangreichen Bibliographie habe ich diesen Themenkreis nicht berücksichtigt.

Arbeiten von Angehörigen der TU Clausthal zu diesem Themenkreis:

Klotz, W.: Generalization of some theorems on sets of multiples and primitive sequences. Acta Arith. 32 (1977), 15–26.

Lucht, L.G.: Dichtentheoretische Sätze der Zahlentheorie. J. Reine Angew. Math. 243 (1970), 192–218 und Band 244 (1970), 220.

Ich selber arbeite daran, interessante Muster in der Menge der Primzahlen nachzuweisen. Dazu gibt es z.B. die Arbeit Elsholtz, C.: Triples of primes in arithmetic progressions (2001).

Die Ausarbeitung stützt sich erheblich auf das Buch Ramseytheorie von Graham et. al. Für die Vorträge stand mir leider nur die erste Auflage zur Verfügung. In der zweiten Auflage befindet sich die Verbesserung von Shelah zu den verbesserten van der Waerden Zahlen.

Ich danke Prof. Hilgert, Prof. Lucht und Prof. Sander und dem Oberseminar Zahlentheorie für die Rückfragen, die zu einer Verbesserung der Darstellung führten.

Literatur

- [1] Ajtai, M.; Szemerédi, E.: Sets of lattice points that form no squares. *Stud. Sci. Math. Hungar.* 9 (1974), 9–11 (1975).
- [2] Alon, N.; Spencer, J. H.: *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1992. xvi+254 pp.
- [3] Anderson, P.G.: A generalization of Baudet’s conjecture (van der Waerden’s theorem). *Amer. Math. Monthly* 83 (1976), 359–361.
- [4] Behrend, F.A.: On sets of integers which contain no three in arithmetic progression, *Proc. Nat. Acad. Sci.* 23 (1946), 331–332.
- [5] V. Bergelson, A. Leibman, Polynomial extensions of van der Waerden’s and Szemerédi’s theorems, *J. Amer. Math. Soc.* 9 (1996), 725–753.
- [6] Berlekamp, E. R.: A construction for partitions which avoid long arithmetic progressions. *Canad. Math. Bull.* 11 1968 409–414.
- [7] Bollobás, B.: *Modern graph theory*. Graduate Texts in Mathematics, 184. Springer-Verlag, New York, 1998. xiv+394 pp.
- [8] Bourgain, J.: On triples in arithmetic progression. *Geom. Funct. Anal.* 9 (1999), no. 5, 968–984.
- [9] Burkill, H.; Mirsky, L.: Monotonicity. *J. Math. Anal. Appl.* 41 (1973), 391–410.
- [10] Deuber, W.: On van der Waerden’s theorem on arithmetic progressions. *J. Combin. Theory Ser. A* 32 (1982), no. 1, 115–118.
- [11] Deuber, W. und Voigt, B.: Der Satz von van der Waerden über arithmetische Progressionen, *Jahresber. Deutsch. Math.-Verein.* 85 (1983), 66–85.
- [12] Erdős, P.; Graham, R. L.: Old and new problems and results in combinatorial number theory: van der Waerden’s theorem and related topics. *Enseign. Math. (2)* 25 (1979), no. 3-4, 325–344 (1980).
- [13] Erdős, P.; Graham, R. L.: Old and new problems and results in combinatorial number theory. *Monographies de L’Enseignement Mathématique*, 28. Université de Genève, 1980. 128 pp.
- [14] Erdős, P.; Turán, P.: On some sequences of integers, *J. London Math. Soc.* 11 (1936), 261–264.
- [15] Fürstenberg, H.: Ergodic behaviour of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. Analyse Math.* 31 (1977), 204–256.
- [16] Fürstenberg, H.; Katznelson, Y.; Ornstein, D.: The ergodic theoretical proof of Szemerédi’s theorem. *Bull. Amer. Math. Soc. (N.S.)* 7 (1982), no. 3, 527–552.
- [17] Gerver, J. L.: The sum of the reciprocals of a set of integers with no arithmetic progression of k terms. *Proc. Amer. Math. Soc.* 62 (1977), no. 2, 211–214.

- [18] Gerver, Joseph L.; Ramsey, L.: Thomas Sets of integers with nonlong arithmetic progressions generated by the greedy algorithm. *Math. Comp.* 33 (1979), no. 148, 1353–1359.
- [19] Gowers, W. T.: Lower bounds of tower type for Szemerédi’s uniformity lemma. *Geom. Funct. Anal.* 7 (1997), no. 2, 322–337.
- [20] Gowers, W. T.: Fourier analysis and Szemerédi’s theorem. *Proceedings of the International Congress of Mathematicians, Vol. I (Berlin, 1998)*. *Doc. Math.* 1998, Extra Vol. I.
- [21] Gowers, W. T.: A new proof of Szemerédi’s theorem for arithmetic progressions of length four. *Geom. Funct. Anal.* 8 (1998), no. 3, 529–551.
- [22] Gowers, W. T.: Rough structure and classification. *GAF A 2000 (Tel Aviv, 1999)*. *Geom. Funct. Anal.* 2000, Special Volume, Part I, 79–117.
- [23] Gowers, W. T.: A new proof of Szemerédi’s theorem. *Geom. Funct. Anal.* 11 (2001), no. 3, 465–588.
- [24] Graham, R. L.: Arithmetic progressions: from Hilbert to Shelah. A joint AMS-MAA lecture presented in Phoenix, Arizona, January 1989. *AMS-MAA Joint Lecture Series*. American Mathematical Society, Providence, RI, 1989. 1 videocassette (NTSC; 1/2 inch; VHS) (60 min.)
- [25] Graham, R. L.; Rothschild, B. L.: A short proof of van der Waerden’s theorem on arithmetic progressions. *Proc. Amer. Math. Soc.* 42 (1974), 385–386.
- [26] Graham, R. L.; Rothschild, Bruce L.; Spencer, Joel H.: *Ramsey theory*. Wiley-Interscience Series in Discrete Mathematics. New York, 1980. (2. Auflage 1990).
- [27] Heath-Brown, D.R.: Integer sets containing no arithmetic progressions, *J. London Math. Soc.* (2) 35 (1987), 385–394.
- [28] Nešetřil, J.: *Ramsey theory*. *Handbook of combinatorics*, Vol. 1, 2, 1331–1403, Elsevier, Amsterdam, 1995.
- [29] Rankin, R.A.: Sets of integers containing not more than a given number of terms in arithmetical progression *Proc. R. Soc. Edinb., Sect. A* 65 (1965), 332–344.
- [30] Riddell, J. A lattice point problem related to sets containing no l -term arithmetic progression. *Canad. Math. Bull.* 14 (1971), 535–538.
- [31] Roth, K.F.: Sur quelques ensembles d’entiers. (French) *C. R. Acad. Sci. Paris* 234, (1952). 388–390.
- [32] Roth, K. F.: On certain sets of integers. *J. London Math. Soc.* 28, (1953). 104–109.
- [33] Roth, K. F.: On certain sets of integers. II. *J. London Math. Soc.* 29, (1954). 20–26.
- [34] Roth, K. F.: Irregularities of sequences relative to arithmetic progressions. IV. *Period. Math. Hungar.* 2 (1972), 301–326.
- [35] Shelah, S.: Primitive recursive bounds for van der Waerden numbers, *J. Amer. Math. Soc.* 1 (1988), 683–697.

- [36] Soifer, A.: Pierre Joseph Henry Baudet: Ramsey theory before Ramsey. *Geombinatorics* 6 (1996), no. 2, 60–70.
- [37] Szemerédi, E.: On sets of integers containing no four elements in arithmetic progression. *Acta Math. Acad. Sci. Hungar.* 20, (1969) 89–104.
- [38] Szemerédi, E.: On sets of integers containing no four elements in arithmetic progression. 1970 Number Theory (Colloq., János Bolyai Math. Soc., Debrecen, 1968) pp. 197–204 North-Holland, Amsterdam
- [39] Szemerédi, E.: On sets of integers containing no k elements in arithmetic progression. *Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974)*, Vol. 2, pp. 503–505. *Canad. Math. Congress*, Montreal, Que., 1975.
- [40] Szemerédi, E.: The number of squares in an arithmetic progression. *Studia Sci. Math. Hungar.* 9 (1974), no. 3–4, 417 (1975).
- [41] Szemerédi, E.: On sets of integers containing no k elements in arithmetic progression. *Acta Arith.* 27 (1975), 199–24.
- [42] Szemerédi, E.: Integer sets containing no arithmetic progressions. *Acta Math. Hungar.* 56 (1990), no. 1–2, 155–158.
- [43] Thouvenot, Jean-Paul: La démonstration de Furstenberg du théorème de Szemerédi sur les progressions arithmétiques. (French) *Séminaire Bourbaki*, 30e année (1977/78), Exp. No. 518, pp. 221–232, *Lecture Notes in Math.*, 710, Springer, Berlin, 1979.
- [44] Vaughan, R.C.: *The Hardy-Littlewood Method* (2nd ed.), *Cambridge Tracts in Mathematics* 125, CUP 1997.
- [45] van der Waerden, B.L.: Beweis einer Baudetschen Vermutung. *Nieuw Arch. Wisk.* 15 (1928), 212–216.
- [46] van der Waerden, B.L.: Wie der Beweis der Vermutung von Baudet gefunden wurde. *Elemente der Math.* 9 (1954), 49–56; Reprint *Elem. Math.* 53 (1998), no. 4, 139–148.
- [47] van der Waerden, B. L.: Wie der Beweis der Vermutung von Baudet gefunden wurde. *Abh. Math. Sem. Univ. Hamburg* 28 (1965), 6–15. Englische Übersetzung in *Studies in Pure Mathematics (Presented to Richard Rado)*, 251–260, Academic Press, London, 1971.
- [48] Wagstaff, S. S.: Some questions about arithmetic progressions. *Amer. Math. Monthly* 86 (1979), no. 7, 579–582.
- [49] Walters, M.: Combinatorial proofs of the polynomial van der Waerden theorem and the polynomial Hales-Jewett theorem. (English. English summary) *J. London Math. Soc.* (2) 61 (2000), no. 1, 1–12.