Chapter 2
From Hilbert's Cube Lemma to Rado's Thesis

Quite a while before Ramsey proved his partition theorem for finite sets some results have been established which can be viewed as the earliest roots of Ramsey theory. The probably first one is due to David Hilbert (1892). In connection with investigations on the irreducibility of rational functions with integer coefficients he proved that for every coloring of some sufficiently large interval $[1, n]$ with r colors, there exist positive integers $a, a_0, \ldots, a_{m-1} \leq n$ such that the affine m-cube

$$\{a + \sum_{i<m} \epsilon_i a_i \mid \epsilon_i \in \{0, 1\} \text{ for every } i < m\}$$

is completely contained in one color class. Apparently neither Hilbert himself nor some other mathematician at that time examined the underlying combinatorial principles of this lemma.

Others happened to a lemma proved by Issai Schur some 25 years later. In reproving a theorem of Dickson on a modular version of Fermat's conjecture, Schur (1916) showed that for every r-coloring of some sufficiently large interval $[1, n]$ there exist positive integers $a_0, a_1 \leq n$ such that the projective 2-cube

$$\{\sum_{i<2} \epsilon_i a_i \mid \epsilon_i \in \{0, 1\} \text{ for every } i < 2\} \backslash \{0\}$$

is completely contained in one color class.

A conjecture of Schur concerning the distribution of quadratic residues, respectively nonresidues modulo p led Schur to a question on arithmetic progressions, which became famous as Baudet's conjecture (cf. Brauer 1973). The problem was solved by Bartel Leendart van der Waerden (1927). The corresponding theorem, well known as van der Waerden's theorem on arithmetic progressions, soon attracted many mathematicians. For example, Khinchin (1952) writing an elementary book on number theoretic problems selected this result as one of his *Three Pearls in Number Theory*.

H.J. Prömel, *Ramsey Theory for Discrete Structures*,
DOI 10.1007/978-3-319-01315-2_2,
© Springer International Publishing Switzerland 2013

Brauer (1928) used van der Waerden's theorem on arithmetic progressions to resolve Schur's conjecture on quadratic residues. In fact, Schur himself suggested a strengthening of van der Waerden's theorem which is in a sense a common generalization of Schur's lemma on projective 2-cubes and van der Waerden's theorem and allows to derive a stronger form of Schur's conjecture (Brauer 1928).

A first culmination point of Ramsey theory was obtained with the work of a student of Schur: Richard Rado. In a series of beautiful papers (Rado 1933a,b, 1943) based on his doctoral dissertation he extended the results of Hilbert, Schur and van der Waerden in a remarkable way. He gave among other results a complete characterization of all systems of homogeneous linear equations $\mathcal{L} = \mathcal{L}(x_0, \ldots, x_{m-1})$ over $\mathbb{Z}$ having the property that for every coloring of $\mathbb{Z}^+$ with finitely many colors, $\mathcal{L}$ has a monochromatic solution. Observe that Schur's lemma essentially says that $x + y = z$ has this property.

One convention: To avoid trivial cases we dismiss throughout this section the number 0. We consider colorings of $[1, n] = \{1, \ldots, n\}$ rather than $n = \{0, \ldots, n-1\}$, and of $\mathbb{N}$, the set of positive integers, instead of ω.

2.1 Hilbert's Cube Lemma

Let a, m and $a_0, \ldots, a_{m-1}$ be positive integers. Then the set

$$\{a + \sum_{i<m} \epsilon_i a_i \mid \epsilon_i \in \{0, 1\} \text{ for every } i < m\}$$

is the affine m-cube generated by $a, a_0, \ldots, a_{m-1}$. Hilbert (1892) proved the following result:

Theorem 2.1 (Hilbert's cube lemma). *Let m and r be positive integers. Then for every r-coloring $\Delta : \mathbb{N} \to r$ of the positive integers there exists an affine m-cube which is monochromatic.*

Hilbert's cube lemma is probably the earliest result which can be viewed as a partition theorem (besides the pigeonhole principle, of course). It was established some 35 years before Ramsey's theorem. Hilbert's proof is written in the style of the late ninetieth century: detailed discussions appealing to the readers mathematical intuition. But despite its unusualness for todays reader the proof is convincing by its clarity and worth reading. So we think it is worth while to include the original proof of Hilbert (though in German). Later in this chapter (Sect. 2.3) we obtain Hilbert's lemma also from van der Waerden's theorem on arithmetic progressions.

Unsere Entwickelungen beruhen auf folgendem Hülfsatze:

Es sei eine unendliche Zahlenreihe $a_1, a_2, a_3, \ldots$ vorgelegt, in welcher allgemein a_s eine der a ganzen positiven Zahlen $1, 2, \ldots, a$ bedeutet; es sei überdies m irgend eine ganze positive Zahl. Dann lassen sich stets m ganze positive Zahlen $\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(m)}$ so bestimmen, dass die 2^m Elemente

$$
\begin{aligned}
&a_\mu,\\
&a_{\mu+\mu^{(1)}},\\
&a_{\mu+\mu^{(2)}}, a_{\mu+\mu^{(1)}+\mu^{(2)}},\\
&a_{\mu+\mu^{(3)}}, a_{\mu+\mu^{(1)}+\mu^{(3)}}, a_{\mu+\mu^{(2)}+\mu^{(3)}}, a_{\mu+\mu^{(1)}+\mu^{(2)}+\mu^{(3)}},\\
&\cdots\cdots\cdots\\
&a_{\mu+\mu^{(\cdot\cdot)}}, a_{\mu+\mu^{(1)}+\mu^{(m)}}, a_{\mu+\mu^{(2)}+\mu^{(\cdot\iota)}}, \ldots, a_{\mu+\mu^{(1)}+\mu^{(2)}+\ldots+\mu^{(m)}},
\end{aligned}
$$

für unendlich viele ganzzahlige Werthe μ sämtlich gleich der nämlichen Zahl $\mathcal{G}$ sind, wo $\mathcal{G}$ eine der Zahlen $1, 2, \ldots, a$ bedeutet. Dabei wird der Index $\mu + \mu^{(1)}$ des zweiten Elementes erhalten, indem man die Zahl $\mu^{(1)}$ zu dem Index μ des ersten Elementes addirt; die Indices des dritten und vierten Elementes entstehen aus den Indices des ersten und zweiten Elementes, indem man zu diesen die Zahl $\mu^{(2)}$ addirt; die Indices des fünften, sechsten, siebenten, achten Elementes entstehen aus den Indices der vier ersten Elemente, wenn man zu diesen die Zahl $\mu^{(3)}$ addirt, und schliesslich erhält man die Indices der 2^{m-1} letzten Elemente, indem man zu den schon bestimmten Indices der 2^{m-1} ersten Elemente die Zahl $\mu^{(m)}$ addiert.

Beim Beweise ist es nothwendig, einzelne Theile der vorgelegten Reihe für sich zu betrachten. Wenn insbesondere i auf einander folgende Elemente der Reihe herausgegriffen werden, etwa die Elemente $a_\mu, a_{\mu+1}, a_{\mu+2}, \ldots, a_{\mu+i-1}$, so nenne ich diese i Elemente ein Intervall der Reihe von der Länge i. Wir grenzen nun innerhalb der vorgelegten Reihe irgend ein Intervall von der Länge $a + 1$ ab. In diesem Intervalle tritt dann mindestens eine der Zahlen $1, 2, \ldots, a$ etwa die Zahl $\mathcal{G}$, zweimal auf, d.h. in dem Intervalle von der Länge $a + 1$ kommt jedenfalls eine der folgenden Gruppirungen vor:

$$
\begin{aligned}
\mathcal{G}_2^{(1)} &= \mathcal{G}\mathcal{G},\\
\mathcal{G}_3^{(1)} &= \mathcal{G}\cdot\mathcal{G},\\
\mathcal{G}_4^{(1)} &= \mathcal{G}\cdot\cdot\mathcal{G},\\
&\cdots\cdots\cdots\cdots\\
\mathcal{G}_{a+1}^{(1)} &= \mathcal{G}\cdots\cdots\cdots\mathcal{G}.
\end{aligned}
$$

Wie schon durch die Schreibweise kenntlich gemacht ist, bedeutet hierin allgemein $\mathcal{G}_s^{(1)}$ ein Intervall von der Länge s, dessen erstes und letztes Element einander gleich, nämlich gleich der Zahl $\mathcal{G}$ sind. Man sieht, dass die Anzahl aller möglichen von einander verschiedenen Gruppirungen $\mathcal{G}_s^{(1)}$ gleich a^2 und somit jedenfalls kleiner als die Zahl $(a + 1)^2$ ist. Wir grenzen jetzt innerhalb der vorgelegten Reihe hinter einander $(a + 1)^2$ Intervalle ab, deren jedes die Länge $a + 1$ besitzt, und betrachten dann das so entstehende Gesammtintervall von der Länge $(a + 1)^3$. In demselben tritt nothwendig mindestens eine der Gruppirungen $\mathcal{G}_s^{(1)}$, etwa die Gruppirung $\mathcal{G}_{\nu^{(1)}}^{(1)}$, zweimal auf, d.h. in dem Intervalle von der Länge $(a + 1)^3$ kommt jedenfalls eine der folgenden Gruppirungen vor:

$$
\begin{aligned}
\mathcal{G}_{2\nu^{(1)}}^{(2)} &= \mathcal{G}_{\nu^{(1)}}^{(1)}\mathcal{G}_{\nu^{(1)}}^{(1)},\\
\mathcal{G}_{2\nu^{(1)}+1}^{(2)} &= \mathcal{G}_{\nu^{(1)}}^{(1)}\cdot\mathcal{G}_{\nu^{(1)}}^{(1)},\\
\mathcal{G}_{2\nu^{(1)}+2}^{(2)} &= \mathcal{G}_{\nu^{(1)}}^{(1)}\cdot\cdot\mathcal{G}_{\nu^{(1)}}^{(1)},\\
&\cdots\cdots\cdots\cdots\\
\mathcal{G}_{(a+1)^3}^{(2)} &= \mathcal{G}_{\nu^{(1)}}^{(1)}\cdots\cdots\cdots\mathcal{G}_{\nu^{(1)}}^{(1)}.
\end{aligned}
$$

Hier bedeutet allgemein $\mathcal{G}_s^{(2)}$ ein Intervall von der Länge s, welches mit der Gruppirung $\mathcal{G}_{\nu^{(1)}}^{(1)}$ beginnt und mit der nämlichen Gruppirung schliesst. Die Anzahl aller von einander verschiedenen Gruppirungen $\mathcal{G}^{(2)}$ ist offenbar kleiner als das Product der Intervalllänge $(a+1)^s$ in die Anzahl aller möglichen Gruppirungen $\mathcal{G}^{(1)}$, und folglich ist jene Anzahl der Gruppirungen $\mathcal{G}_s^{(2)}$ jedenfalls kleiner als $(a+1)^5$. Wenn wir daher innerhalb der vorgelegten Reihe hinter einander $(a+1)^5$ Intervalle abgrenzen und zwar ein jedes von der Länge $(a+1)^3$, so tritt in dem so entstehenden Intervalle von der Gesammtlänge $(a+1)^8$ mindestens eine der Gruppirungen $\mathcal{G}_s^{(2)}$, etwa die Gruppirung $\mathcal{G}_{\nu(2)}^{(2)}$, zweimal auf, d. h. in dem Intervalle von der Länge $(a+1)^8$ kommt jedenfalls eine der folgenden Gruppirungen vor:

$$\begin{aligned}
\mathcal{G}_{2\nu^{(2)}}^{(3)} &= \mathcal{G}_{\nu^{(2)}}^{(2)}\mathcal{G}_{\nu^{(2)}}^{(2)},\\
\mathcal{G}_{2\nu^{(2)}+1}^{(3)} &= \mathcal{G}_{\nu^{(2)}}^{(2)}\cdot\mathcal{G}_{\nu^{(2)}}^{(2)},\\
\mathcal{G}_{2\nu^{(2)}+2}^{(3)} &= \mathcal{G}_{\nu^{(2)}}^{(2)}\cdot\cdot\mathcal{G}_{\nu^{(2)}}^{(2)},\\
&\cdots\cdots\cdots\cdots\\
\mathcal{G}_{(a+1)^8}^{(3)} &= \mathcal{G}_{\nu^{(2)}}^{(2)}\cdots\cdots\cdots\mathcal{G}_{\nu^{(2)}}^{(2)}.
\end{aligned}$$

Hier bedeutet allgemein $\mathcal{G}_s^{(3)}$ ein Intervall von der Länge s, welches mit der Gruppirung $\mathcal{G}_{\nu(2)}^{(2)}$ beginnt und mit der nämlichen Gruppirung schliesst.

Nach m-maliger Anwendung des nämlichen Verfahrens gelangen wir zu Gruppirungen von der Gestalt:

$$\mathcal{G}^{(m)} = \mathcal{G}^{(m-1)}\cdots\cdots\cdots\mathcal{G}^{(m-1)}$$

und erkennen, dass in jedem Intervall der Reihe von einer gewissen Länge ℓ nothwendig eine jener Gruppirungen $\mathcal{G}^{(m)}$ vorkommen muss. Dabei bedeutet ℓ eine bestimmte endliche und nur von a und m abhängige Zahl. Die Anzahl aller von einander verschiedenen Gruppirungen $\mathcal{G}^{(m)}$ ergiebt sich wiederum kleiner als eine gewisse endliche Zahl k welche leicht aus a und m berechnet werden kann. In der vorgelegten Reihe können wir nun hinter einander beliebig viele Intervalle von der Länge abgrenzen, und es folgt daher, dass es unter den Gruppirungen $\mathcal{G}^{(m)}$ nothwendig eine giebt, welche in der vorgelegten Reihe unendlich oft vorkommt. Diese Gruppirung sei die folgende

$$\mathcal{G}_{\nu^{(m)}}^{(m)} = \mathcal{G}_{\nu^{(m-1)}}^{(m-1)}\cdots\cdots\cdots\mathcal{G}_{\nu^{(m-1)}}^{(m-1)},$$

wo $\mathcal{G}_{\nu^{(m)}}^{(m)}$ und $\mathcal{G}_{\nu^{(m-1)}}^{(m-1)}$, Intervalle von der Länge $\nu^{(m)}$ beziehungsweise von der Länge $\nu^{(m-1)}$ bedeuten.

Wir erkennen hieraus leicht die Richtigkeit des obigen Hülfsatzes. Es ist nämlich die Gruppirung $\mathcal{G}_{\nu^{(m)}}^{(m)}$ durch die folgenden Recursionsformeln bestimmt:

$$\begin{aligned}
\mathcal{G}_{\nu^{(1)}}^{(1)} &= \mathcal{G}\cdots\cdots\cdots\mathcal{G},\\
\mathcal{G}_{\nu^{(2)}}^{(2)} &= \mathcal{G}_{\nu^{(1)}}^{(1)}\cdots\cdots\cdots\mathcal{G}_{\nu^{(1)}}^{(1)},\\
\mathcal{G}_{\nu^{(3)}}^{(3)} &= \mathcal{G}_{\nu^{(2)}}^{(2)}\cdots\cdots\cdots\mathcal{G}_{\nu^{(2)}}^{(2)},\\
&\cdots\cdots\cdots\cdots\\
\mathcal{G}_{\nu^{(m)}}^{(m)} &= \mathcal{G}_{\nu^{(m-1)}}^{(m-1)}\cdots\cdots\cdots\mathcal{G}_{\nu^{(m-1)}}^{(m-1)}.
\end{aligned}$$

wo stets die unteren Indices die Anzahl der Elemente angeben, aus denen die betreffenden Intervalle bestehen. Ich setze

$$\begin{aligned}
\mu^{(1)} &= \nu^{(1)} - 1, \\
\mu^{(2)} &= \nu^{(2)} - \nu^{(1)}, \\
\mu^{(3)} &= \nu^{(3)} - \nu^{(2)}, \\
&\cdots\cdots\cdots\cdots \\
\mu^{(m)} &= \nu^{(m)} - \nu^{(m-1)},
\end{aligned}$$

und behaupte dann, dass die so entstehenden, ganzen positiven Zahlen $\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(m)}$ von derjenigen Beschaffenheit sind, welche unser Hülfsatz verlangt. In der That: es ist eben bewiesen worden, dass in der vorgelegten Reihe $a_1, a_2, a_3, \ldots$ die Gruppirung $\mathcal{G}^{(m)}_{\nu^{(m)}}$ unendlich oft vorkommt, d. h. es giebt unendlich viele ganzzahlige Werthe von μ, für welche

$$a_\mu a_{\mu+1} \ldots a_{\mu+\nu^{(m)}-1} = \mathcal{G}^{(m)}_{\nu(m)}$$

wird. Aus dem Aufbau der Gruppirung $\mathcal{G}^{(m)}_{\nu^{(m)}}$ folgt dann

$$\begin{aligned}
a_\mu &= \mathcal{G}, \\
a_{\mu+\mu^{(1)}} &= \mathcal{G}, \\
a_{\mu+\mu^{(2)}} &= a_{\mu+\mu^{(1)}+\mu^{(2)}} = \mathcal{G}, \\
a_{\mu+\mu^{(3)}} &= a_{\mu+\mu^{(1)}+\mu^{(3)}} = a_{\mu+\mu^{(2)}+\mu^{(3)}} = a_{\mu+\mu^{(1)}+\mu^{(2)}+\mu^{(3)}} = \mathcal{G}, \\
&\cdots\cdots\cdots\cdots \\
a_{\mu+\mu^{(m)}} &= a_{\mu+\mu^{(1)}+\mu^{(m)}} = a_{\mu+\mu^{(2)}+\mu^{(m)}} = \ldots = a_{\mu+\mu^{(1)}+\mu^{(2)}+\ldots+\mu^{(m)}} = \mathcal{G},
\end{aligned}$$

und damit ist der Hülfsatz bewiesen.

2.2 Schur's Lemma

Theorem 2.2 (Schur's lemma). *Let r be a positive integer. Then there exists a least positive integer $n = S(r)$, such that for every coloring $\Delta : [1, n] \to r$ there exist positive integers $x, y \leq n$ satisfying*

$$\Delta(x) = \Delta(y) = \Delta(x + y).$$

Moreover, $S(r) \leq e r!$, where e is the base of the natural logarithm.

Let m and $a_0, \ldots, a_{m-1}$ be positive integers. Then the set

$$\{\sum_{i<m} \epsilon_i a_i \mid \epsilon_i \in \{0, 1\} \text{ for every } i < m\} \backslash \{0\}$$

is the projective m-cube generated by $a_0, \ldots, a_{m-1}$. Using this terminology Schur's lemma can be rephrased by saying that for every coloring $\Delta : [1, \lfloor er! \rfloor] \to r$ there exists a projective 2-cube which is monochromatic. Hence the lemma is a projective analogue to Hilbert's (affine) cube lemma for $m = 2$.

Hilbert used his lemma as a tool to obtain certain results on the irreducibility of rational functions with integer coefficients. Schur established his lemma to give an easy proof and moreover to extend a number theoretic theorem of Dickson showing that for each r the congruence

$$x^r + y^r \equiv z^r \pmod{p}$$

has solutions for all sufficiently large primes p.

We give two proofs of Schur's lemma. The first one follows the lines of Schur's original proof yielding the bound $S(r) \leq er!$ The second one is an application of Ramsey's theorem.

First Proof of Schur's lemma. Let $n_0 \geq er!$ and let $\Delta : [1, n_0] \to r$ be an r-coloring of the first n_0 positive integers. Assume that there do not exist integers $x, y \leq n_0$ such that $\Delta(x) = \Delta(y) = \Delta(x + y)$.

Let $i_0 < r$ be the color which occurs most frequently under the n_0 elements and let $\Delta^{-1}(i_0) = \{x_0, \ldots, x_{n_1-1}\}$ be in ascending order. Observe that $n_0 \leq rn_1$.

Consider $N_0 = \{x_i - x_0 \mid 1 \leq i < n_1\}$. By assumption $N_0 \cap \Delta^{-1}(i_0) = \emptyset$. Let i_1 be the most frequent color under the elements of N_0 and let $N_0 \cap \Delta^{-1}(i_1) = \{y_0, \ldots, y_{n_2-1}\}$ be in ascending order. Observe that $n_1 - 1 \leq (r-1)n_2$.

Consider $N_1 = \{y_i - y_0 \mid 1 \leq i < n_2\}$. By assumption $N_1 \cap \Delta^{-1}(i_0) = \emptyset$ and $N_1 \cap \Delta^{-1}(i_1) = \emptyset$. Let i_2 be the most frequent color under the elements of N_1 and let $N_1 \cap \Delta^{-1}(i_2) = \{z_0, \ldots, z_{n_3-1}\}$ be in ascending order. Observe that $n_2 - 1 \leq (r-2)n_3$.

Continue this procedure until some n_j becomes 1. At latest $n_r = 1$, as otherwise N_r contains two elements whose difference cannot not be colored by any of the r colors.

Inserting the above inequalities into each other gives eventually

$$n_0 \leq r! \cdot \left(1 + \sum_{i=1}^{r-1} \frac{1}{i!}\right) < r! \cdot e,$$

a contradiction to the choice of n_0. Hence, there exist x, y such that $\Delta(x) = \Delta(y) = \Delta(x + y)$. □

Second proof of Schur's lemma. Let n be according to the finite Ramsey theorem such that $n \to (3)^2_r$ and let an r-coloring $\Delta : [1, n] \to r$ be given. This induces an r-coloring $\Delta^* : [n]^2 \to r$ by $\Delta^*(a, b) = \Delta(b - a)$ for $a < b$. By choice of n there exist $0 \leq u < v < w < n$ so that

$$\Delta^*(u, v) = \Delta^*(v, w) = \Delta^*(u, w)$$

and, hence, $\Delta(v-u) = \Delta(w-v) = \Delta(w-u)$. Putting $x = v-u$ and $y = w-v$ proves Schur's lemma. □

Irving (1973) has slightly improved Schur's upper bound on $S(r)$ from $\lfloor r!e \rfloor$ to $\lfloor r!(e - \frac{1}{24}) \rfloor$. A lower bound is given in Fredricksen (1975, 1979), viz. $S(r) \geq c(315)^{\frac{r}{5}}$ for an appropriate constant c, cf. also Sect. 7.5.

2.3 Van der Waerden's Theorem

Schur, working on the distribution of quadratic residues and nonresidues, conjectured that for every k and every sufficiently large prime p there exist k consecutive numbers which are quadratic residues as well as k consecutive numbers which are quadratic nonresidues modulo p. To attack this conjecture he tried first to prove that for every k there exists n so that for every 2-coloring of $1, \ldots, n$ one of the two color classes contains an arithmetic progression of length k. He failed and both questions remained open for several years (cf. Brauer 1973).

Van der Waerden learned about the conjecture on arithmetic progressions most probably from P.J.H. Baudet at that time a young Dutch student in Göttingen. So his answer to this conjecture (van der Waerden 1927) is entitled *Beweis einer Baudetschen Vermutung*.

Theorem 2.3 (van der Waerden). *Let k and r be positive integers. Then there exists a least positive integer $n = W(k,r)$ such that for every r-coloring $\Delta : [1,n] \to r$ there exists an arithmetic progression*

$$\{a + i\,d \mid i < k\} \subseteq [1,n]$$

of length k which is monochromatic.

Years later, van der Waerden (1954, 1971) gave a personal account on *How the proof of Baudet's conjecture was found* – by now a classical contribution to the psychology of invention in mathematics.

Proof of Theorem 2.3. We prove actually something stronger than van der Waerden's theorem, namely:

Let k, m and r be positive integers. Then there exists a least positive integer $n = S(k,m,r)$ such that for every coloring $\Delta : [1,n] \to r$ there exist positive integers a and $d_0, \ldots, d_{m-1}$ so that $a + k \cdot \sum_{i<m} d_i \leq n$ and

$$\Delta(a + \textstyle\sum_{i<m} g_i d_i) = \Delta(a + \sum_{i<m} h_i d_i)$$

whenever $\boldsymbol{g}, \boldsymbol{h} \in (k+1)^m$, where $\boldsymbol{g} = (g_0, \ldots, g_{m-1})$ and $\boldsymbol{h} = (h_0, \ldots, h_{m-1})$, agree up to their last occurrence of k (in $\boldsymbol{g}$ or $\boldsymbol{h}$). Note: this implies that any combination of $\boldsymbol{g}, \boldsymbol{h} \in k^m$ is allowed, as then neither of them contains any k.

A set $\{a + \sum_{i<m} g_i d_i \mid \boldsymbol{g} \in k^m\}$ is called *m-fold arithmetic progression*. Observe that for $m = 1$ we get the standard arithmetic progression of length k, thus $W(k,r) \leq S(k,1,r)$ and van der Waerden's theorem follows.

We show the following two inequalities hold for all k, m, r:

1. $S(k, m+1, r) \leq S(k, m, r) \cdot S(k, 1, r^{S(k,m,r)})$,
2. $S(k+1, 1, r) \leq S(k, r, r)$

Together with the trivial observation that $S(1,1,r) = 2$ for every r these inequalities immediately yield the proof of the statement by induction on m and k.

Proof of (1): Let $M = S(k,m,r)$ and $N = S(k,1,r^{S(k,m,r)})$ and consider $\Delta : [1, M \cdot N] \to r$. This induces a coloring $\Delta_N : [1,N] \to r^M$ by

$$\Delta_N(x) = \langle \Delta((x-1)M + j) \mid 1 \leq j \leq M \rangle.$$

By choice of N there exist positive integers b and d so that $\{b + jd \mid j < k\} \subseteq [1,N]$ and $\Delta_N \restriction \{b + jd \mid j < k\}$ is a constant coloring. Observe that this means that for any $1 \leq j \leq M$ we have

$$\Delta((b\text{–}1)M + j) = \Delta((b\text{–}1+d)M + j) = \ldots = \Delta((b\text{–}1+(k\text{–}1)d)M + j). \quad (2.1)$$

Next consider $\Delta_M : [(b-1)M + 1, bM] \to r$ where $\Delta_M = \Delta \restriction [(b-1)M + 1, bM]$. By choice of M there exist positive integers $a, d_0, \ldots, d_{m-1}$ so that the m-fold arithmetic progression $\{a + \sum_{i<m} g_i d_i \mid \boldsymbol{g} \in (k+1)^m\}$ is completely contained in $[(b-1)M + 1, bM]$ and

$$\Delta(a + \textstyle\sum_{i<m} g_i d_i) = \Delta(a + \sum_{i<m} h_i d_i) \quad (2.2)$$

for all $\boldsymbol{g}, \boldsymbol{h} \in (k+1)^m$ which agree up to their last occurrence of k. Let $d_m := dM$. We claim that then

$$\Delta(a + \textstyle\sum_{i\leq m} g_i d_i) = \Delta(a + \sum_{i\leq m} h_i d_i)$$

for all $\boldsymbol{g}, \boldsymbol{h} \in (k+1)^{m+1}$ which agree up to their last occurrence of k. Note that the proof of this claim implies that (1) holds. In order to see why the claim holds observe first that if $g_m = k$ or $h_m = k$ then $\boldsymbol{g} = \boldsymbol{h}$ and the claim holds trivially. So assume $g_m, h_m < k$. Then the choice of $d_m = dM$ and (2.1) implies that

$$\Delta(a + \textstyle\sum_{i\leq m} g_i d_i) = \Delta(a + \sum_{i<m} g_i d_i)$$

and

$$\Delta(a + \sum_{i\leq m} h_i d_i) = \Delta(a + \sum_{i<m} h_i d_i).$$

The claim thus follows immediately from (2.1).

Proof of (2): Let $N = S(k, r, r)$ and consider $\Delta : [1, N] \to r$. Then there exist $a, d_0, \ldots, d_{r-1}$ such that

$$\Delta(a + \textstyle\sum_{i<r} g_i d_i) = \Delta(a + \sum_{i<r} h_i d_i), \tag{2.3}$$

whenever $\boldsymbol{g}, \boldsymbol{h} \in (k+1)^r$ agree up to their last occurrence of k. Consider

$$\begin{aligned} \boldsymbol{g}^0 &= (\,0\,,\ 0\,,\ldots,\ 0\,)\\ \boldsymbol{g}^1 &= (\,k\,,\ 0\,,\ldots,\ 0\,)\\ &\ \ \vdots\\ \boldsymbol{g}^r &= (\,k\,,\ k\,,\ldots,\ k\,). \end{aligned}$$

By the pigeonhole principle two of these words, say $\boldsymbol{g}^\mu$ and $\boldsymbol{g}^\nu$ where $\mu < \nu$, have the property that $\Delta(a + \sum_{i<r} g_i^\mu d_i) = \Delta(a + \sum_{i<r} g_i^\nu d_i)$. More precisely,

$$\Delta(a + k\textstyle\sum_{i<\mu} d_i) = \Delta(a + k\sum_{i<\mu} d_i + k\sum_{i=\mu}^{\nu-1} d_i).$$

On the other hand, from (2.3) we have

$$\Delta(a + k\textstyle\sum_{i<\mu} d_i) = \Delta(a + k\sum_{i<\mu} d_i + j\sum_{i=\mu}^{\nu-1} d_i),$$

for every $j < k$. Thus, setting $a' = a + k\sum_{i<\mu} d_i$ and $d = \sum_{i=\mu}^{\nu-1} d_i$, we have that

$$\Delta\restriction\{a' + j \cdot d \mid j < k+1\}$$

is a constant coloring. Thus (2) holds. □

The proof given above follows Graham and Rothschild (1974). Like the original proof of van der Waerden's theorem this proof also uses substantially that the assertion is known to be true for $k-1$ and all r in order to derive it for k and some fixed r, say $r = 2$. Combinatorial proofs where the color number is fixed throughout the whole proof were obtained by Deuber (1982) using ideas from the proof of Hales-Jewett's theorem (cf. Sect. 4.1) and by Taylor (1982) giving a combinatorial version of the (Furstenberg and Weiss 1978) topological proof of van der Waerden's theorem.

The above proof has one disadvantage: the fact that it uses some kind of double induction yields an upper bound even on $W(k) := W(k,\ 2)$ which is not primitive recursive. In contrary to this the best lower bound currently available is $W(k+1) \geq k2^k$, for k prime, which is due to Berlekamp (1968). Determining the order of magnitude of $W(k)$ or even proving that $W(k)$ increases slower than the Ackermann function has long been a challenging open problem in Ramsey theory. This was finally solved by Shelah (1988) who proved that the van der Waerden numbers are primitive recursive. The currently best asymptotic upper bound is by Gowers (2001). Some known exact values of $W(k)$ are $W(2) = 3$, $W(3) = 9$, $W(4) = 35$ and $W(5) = 178$ (see Chvátal 1970; Stevens and Shantaram 1978).

Let $H(m,r)$ denote the least integer for which the assertion of Hilbert's cube lemma (Theorem 2.1) is valid, i.e., the smallest number such that for every r-coloring $\Delta : [1, H(m,r)] \to r$ there exists an affine m-cube which is monochromatic with respect to Δ. Obviously, $H(m,r) \leq W(m,r)$, as can be seen as follows. Let $\{a + jd \mid j < m\}$ be a monochromatic arithmetic progression. Then $a, d, d, \ldots, d$ (m many d's) generate a monochromatic affine m-cube proving Hilbert's cube lemma.

But in fact, $H(m,r)$ is much smaller than the van der Waerden number given above. Brown et al. (1985) showed using a result on B_2-sets that $H(2,r)$ is only quadratic in r, more precisely,

$$H(2,r) = (1 + o(1))r^2.$$

Moreover, examining Hilbert's original proof they observed that in general $H(m,r) \leq r^{c^m}$ for an appropriate constant c. In other words, even for arbitrary m, the function $H(m,r)$ is bounded by a polynomial in r.

2.4 Schur's Extension of Van der Waerden's Theorem

"A few days" after van der Waerden answered Schur's question on arithmetic progressions, Brauer (1928) was able to use van der Waerden's result to resolve Schur's conjecture on quadratic residues and nonresidues. But Brauer's paper contains also a strengthening of van der Waerden's theorem (and of Schur's lemma) which he attributes to Schur (cf. also Brauer 1973):

Theorem 2.4. *Let k and r be positive integers. Then there exists a least positive integer $n = SB(k,r)$ such that for every r-coloring $\Delta : [1,n] \to r$ there exists an arithmetic progression*

$$\{a + i\,d \mid i < k\} \subseteq [1,n]$$

of length k which is monochromatic and its difference d is in the same color, i.e., $\Delta](\{a + i\,d \mid i < k\} \cup \{d\})$ is a constant coloring.

Proof. We proceed by induction on the color number r, the case $r = 1$ being trivial for every k.

Assume that the existence of $SB(k, r-1)$ has been established for some $r > 1$. Choose $n = W(k \cdot SB(k, r-1) + 1, r)$ and let $\Delta : [1, n] \to r$ be an arbitrary r-coloring. By choice of n there exists an arithmetic progression

$$\{a + jd' \mid j \leq k \cdot SB(k, r-1)\}$$

which is monochromatic with respect to Δ, say in color $r - 1$.

Now either for some j, $0 < j \leq SB(k, r-1)$, we have $\Delta(jd') = r-1$. In this case we are done with a and $d = jd'$. Or $\Delta\rceil\{jd' \mid 0 < j \leq SB(k, r-1)\}$ is an $(r-1)$-coloring. In that case using the inductive hypothesis finishes the proof. □

We outline the proof of Schur's conjecture using this strengthening of van der Waerden's theorem.

Let p be a prime number and let n be prime to p. Recall that n is a quadratic residue modulo p if $x^2 \equiv n \pmod{p}$ for some positive integer x; otherwise n is a quadratic nonresidue modulo p. Thus the set of integers is divided into three classes, the class of quadratic residues, the class of quadratic nonresidues and the multiples of p. The Legendre symbol $(\frac{n}{p})$ is used to indicate the quadratic character of a number. Its value is ± 1 according to whether n is (or is not) a quadratic residue modulo p. There exist $\frac{1}{2}(p-1)$ quadratic residues, respectively, $\frac{1}{2}(p-1)$ non-residues modulo p in $\mathbb{Z}_p$.

Theorem 2.5. *Let k be a positive integer. Then there exists a positive integer $n = n(k)$ such that for every prime number $p > n$ there exist k consecutive integers which are quadratic residues modulo p and there exist k consecutive integers which are quadratic nonresidues modulo p.*

Proof. First we show that for every sufficiently large prime there exist k consecutive integers which are quadratic residues modulo p.

Let $n = SB(k, 2)$ and $p > n$ be a prime number. Color $[1,\ p-1]$ according to being a quadratic residue modulo p. By choice of n there exists an arithmetic progression

$$\{a + jd \mid j < k\} \subseteq [1,\ p-1]$$

which is monochromatic and its difference d is in the same color, i.e., the Legendre symbol $(\frac{x}{p})$ is constant on $\{a + jd \mid j < k\} \cup \{d\}$.

As the product of two quadratic residues as well as the product of two quadratic nonresidues are quadratic residues, whereas the product of a quadratic residue and a quadratic nonresidue is a nonresidue we deduce that

$$\{\frac{a + jd}{d} \mid j < k\}$$

(with division in the Gallois field $\mathbb{Z}_p$) is a sequence of k consecutive quadratic residues modulo p proving the first part of the theorem.

Now let $\ell = (k! - 1)(k-1) + 1$, $n = SB(\ell, 2)$ and $p > n$ be a prime number. According to the first part of the proof there exists a sequence

$$\{b + j \mid j < \ell\} \subseteq [1, p-1]$$

of ℓ consecutive quadratic residues modulo p. Let d be the smallest nonresidue modulo p. If $d < k!$ then

$$\{\frac{b + jd}{d} \mid j < k\}$$

(with division again in the field $\mathbb{Z}_p$) is a sequence of k consecutive nonresidues. So we can assume that $d \geq k!$. But then $d = k!m + c$ where $c < k!$. Therefore $c - d \equiv 0 \pmod{j}$ and hence $c - d + jd \equiv 0 \pmod{j}$ for every $0 < j < k$. But this gives that

$$\frac{c - d}{j} + d < d$$

and so by assumption $\frac{c-d}{j} + d$ is a quadratic residue. Since $j < k \leq d$, also $(\frac{c-d}{j} + d)j$ is a quadratic residue. Therefore

$$\{(c - d) + jd \mid j < k\}$$

is a progression of quadratic residues but its difference d is a nonresidue. Dividing by d yields the desired result. □

2.5 Rado's Thesis

2.5.1 Partition Regular Systems of Homogenous Linear Equations

Let $A\boldsymbol{x} = \mathbf{0}$ be a system of homogenous linear equations in n variables with integer coefficients. Then $A\boldsymbol{x} = \mathbf{0}$ is *partition regular* if for every coloring of the positive integers with finitely many colors there exists a monochromatic solution, in other words, there exist positive integers $x_0, \ldots, x_{n-1}$ (not necessarily distinct) so that $A(x_0, \ldots, x_{n-1})^T = \mathbf{0}$ and $x_0, \ldots, x_{n-1}$ are all in the same color.

Schur's lemma asserts that

$$x_0 + x_1 - x_2 = 0$$

is partition regular, Schur's extension of van der Waerden's theorem that for every k the system

$$x_1 = x_0 + d$$
$$x_2 = x_1 + d$$

$$\vdots$$

$$x_k = x_{k-1} + d$$

is partition regular and Hilbert's cube lemma implies that

$$a + \sum_{i \in I} x_i = x_I, \quad I \subseteq n,\ I \neq \emptyset$$

is a partition regular system of equations.

Observe that using a compactness argument as in the proof of the finite Ramsey theorem (Theorem 1.2) it follows immediately that if $A\boldsymbol{x} = \boldsymbol{0}$ is partition regular then for each positive integer r there exists already a positive integer $N = N(A, r)$ such that for every r-coloring of $[1, N]$ there exists a monochromatic solution of $A\boldsymbol{x} = \boldsymbol{0}$ in $[1, N]$.

The notion of partition regularity is defined only for positive integers and all examples considered so far deal only with colorings of positive integers. One might think that additional linear systems of equations turn out to be partition regular if we consider r-colorings of nonzero rationals. The following lemma shows that this is not the case.

Lemma 2.6. *Let A be a matrix with integer coefficients. Then the following properties are equivalent:*

(1) *$Ax = 0$ is partition regular,*
(2) *For every coloring of the non-zero integers with finitely many colors there exists a monochromatic solution of $A\boldsymbol{x} = \boldsymbol{0}$,*
(3) *For every coloring of the non-zero rationals with finitely many colors there exists a monochromatic solution of $A\boldsymbol{x} = \boldsymbol{0}$.*

Proof. Since $\mathbb{N} \subseteq \mathbb{Z}\backslash\{0\} \subseteq \mathbb{Q}\backslash\{0\}$, we have trivially the implications from (1) to (2) and from (2) to (3).

Assume (3) and let r be a positive integer. By a compactness argument (König's lemma) there exists a finite set $S \subseteq \mathbb{Q}\backslash\{0\}$ such that for every r-coloring of S there exists a monochromatic solution of $A\boldsymbol{x} = \boldsymbol{0}$ in S. Multiply S with an appropriate integer c such that $\{cs \mid s \in S\} \subseteq \mathbb{Z}\backslash\{0\}$. Then for every r-coloring of $\{cs \mid s \in S\}$ there exists a monochromatic solution of $A\boldsymbol{x} = \boldsymbol{0}$, showing (2).

Now assume (2) and let $\Delta : \mathbb{N} \to r$ be a coloring. Define $\Delta' : \mathbb{Z}\backslash\{0\} \to 2r$ by

$$\Delta'(z) = \begin{cases} \Delta(z) & \text{if } z > 0 \\ \Delta(-z) + r & \text{if } z < 0. \end{cases}$$

Then by homogeneity the required result follows. □

Observe that by homogeneity we could also replace A by a matrix with rational coefficients.

Based on his thesis written under the supervision of Schur, Rado provided a complete characterization of all systems of homogenous linear equations which are partition regular. The crucial notion in this characterization is the column property of a matrix.

Definition 2.7. Let A be a matrix with integer coefficients, say $A = (\boldsymbol{a}^0, \ldots, \boldsymbol{a}^{n-1})$ where the $\boldsymbol{a}^i$ are the columns of A. Then A has the *column property* if there exists a partition of n, say $n = I_0 \cup \ldots \cup I_\ell$ for some $\ell < n$, such that

1. $\sum_{i \in I_0} \boldsymbol{a}^i = 0$, i.e., the columns in I_0 add up to $\mathbf{0}$, and
2. for every $j < \ell$ there exist rational numbers ξ_{ij} such that

$$\sum_{i \in I_{j+1}} \boldsymbol{a}^i = \sum_{i \in \bigcup_{\nu \leq j} I_\nu} \xi_{ij} \boldsymbol{a}^i,$$

i.e., the sum of the columns in I_{j+1} is a rational linear combination of the columns in the previous classes $I_0 \cup \ldots \cup I_j$.

We now consider some examples.

(1) The matrix

$$(1, 1, -1)$$

which describes Schur's equation $x_0 + x_1 = x_2$ obviously has the column property.

(2) The matrix of the system of equations

$$x_{i+1} = x_i + d, \quad i < k$$

has the column property choosing I_0 and I_1 as depicted below:

$$\underbrace{\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}}_{I_1} \underbrace{\begin{pmatrix} 1 & -1 & 0 & \ldots & 0 \\ 0 & 1 & -1 & \ldots & 0 \\ \vdots & & \ddots & & \\ 0 & \ldots & 0 & 1 & -1 \end{pmatrix}}_{I_0}$$

(3) As a third example we consider the matrix corresponding to the system

$$\sum_{i \in I} x_i = x_I, \quad I \subseteq n, I \neq \emptyset,$$

which is a projective version of the system of equations we get from Hilbert's cube lemma. For $n = 2$ we obtain the matrix of Example (1) where the trivial equations are omitted. For $n = 3 = \{0, 1, 2\}$ the corresponding matrix is given below

$$\begin{pmatrix} 1 & 1 & 1 & -1 & 0 & \dots & 0 \\ 1 & 1 & 0 & 0 & -1 & \dots & 0 \\ 1 & 0 & 1 & & & & \\ 1 & 0 & 0 & & & & \\ 0 & 1 & 1 & & & & \\ 0 & 1 & 0 & & & & \\ 0 & 0 & 1 & 0 & \dots & \dots & -1 \end{pmatrix}$$

It can easily be seen that such matrices have the column property: assume that the matrix is arranged so that the rows are ordered lexicographically from the top to the bottom with respect to $1 > 0$ and the ith row contains exactly one -1 positioned in the $(n + i)$th column (compare the picture above). Then

$$I_0 = \{a^0\} \cup \{a^n, \dots, a^{n+2^{n-1}-1}\}, \text{ and}$$

$$I_j = \{a^j\} \cup \{a^{n+\sum_{i=1}^{j} 2^{n-i}}, \dots, a^{n+\sum_{i=1}^{j+1} 2^{n-i}-1}\} \quad \text{for } 1 \leq j < n,$$

gives the desired column partition.

Theorem 2.8 (Rado). *Let A be a matrix with integer coefficients. Then the homogeneous system $A\boldsymbol{x} = \mathbf{0}$ of linear equations is partition regular if and only if the matrix A has the column property.*

We postpone the proof of Rado's theorem until we have introduced the so-called (m, p, c)-sets which can be viewed as generalizations of arithmetic progressions. The notion of (m, p, c)-sets was invented by W. Deuber in his doctoral dissertation where he proved a partition theorem for these sets as a tool to answer a long standing conjecture of Rado in the affirmative (Deuber 1973).

We will use this partition theorem for (m, p, c)-sets to prove Rado's theorem.

2.5.2 (m, p, c)-Sets

Definition 2.9. Let m, p, c be positive integers. A set $M \subseteq \mathbb{N}$ is an (m, p, c)-set if there exist positive integers $x_0, \dots, x_m$ such that

$$M = M_{p,c}(x_0, \dots, x_m)$$

$$= \{cx_i + \sum_{j=i+1}^{m} \xi_j x_j \mid \xi_j \in [-p, p] \cap \mathbb{Z} \text{ for every } j \in [i + 1, m] \text{ and } i \leq m\}.$$

Observe that a $(1,k,1)$-set is an arithmetic progression together with its difference and an $(n,1,1)$-set contains solutions to the system of equations given in Example (3) in the last section. Intuitively speaking, (m,p,c)-sets are m-fold arithmetic progressions together with c-fold differences.

We show that every system $A\boldsymbol{x} = \boldsymbol{0}$ of homogeneous linear equations given by a matrix A having the column property admits to find positive integers m, p and c such that every (m,p,c)-set contains a solution of $A\boldsymbol{x} = \boldsymbol{0}$.

Together with a partition theorem for (m,p,c)-sets this will yield that $A\boldsymbol{x} = \boldsymbol{0}$ is partition regular.

Lemma 2.10. *Let A be a matrix with integer coefficients having the column property. Then there exist positive integers m, p and c such that every (m,p,c)-set contains a solution of $A\boldsymbol{x} = \boldsymbol{0}$.*

Proof. Let $A = (\boldsymbol{a}^0, \dots, \boldsymbol{a}^{n-1})$. By definition there exists a partition $n = I_0 \cup \ldots \cup I_\ell$ such that $\sum_{i\in I_0} \boldsymbol{a}^i = \boldsymbol{0}$ and for every $j < \ell$ there exist rationals ξ_{ij} so that

$$\sum_{i\in I_{j+1}} \boldsymbol{a}^i = \sum_{i\in\bigcup_{v\le j} I_v} \xi_{ij}\boldsymbol{a}^i.$$

Put $m = \ell + 1$ and let c be the least common multiple of the denominators of the ξ_{ij}. Finally, define $\tilde{p}$ to be the maximum of the absolute values of the ξ_{ij} and put $p = \tilde{p}\cdot c$. We claim that m, p and c have the desired properties.

We now show by induction on k that every (k,p,c)-set contains a solution of the matrix A_k consisting of those columns of A belonging to the classes $\bigcup_{i\le k} I_i$.

Clearly, this is true for the matrix $A_0 = (\boldsymbol{a}^i \mid i \in I_0)$ since every singleton provides a solution of $A_0\boldsymbol{x} = \boldsymbol{0}$. Now consider the $(k+1,p,c)$-set $M = M_{p,c}(x_0,\dots,x_{k+1})$ for some $k \ge 0$. By induction hypothesis we know that the (k,p,c)-set $M_{p,c}(x_0,\dots,x_k) \subseteq M$ contains a solution of $A_k = (\boldsymbol{a}^i \mid i \in \bigcup_{j\le k} I_j)$, say

$$\sum_{i\in\bigcup_{j\le k} I_j} y_i\boldsymbol{a}^i = \boldsymbol{0}, \quad \text{where } y_i \in M_{p,c}(x_0,\dots,x_k) \text{ for every } i.$$

By the column property of A and by choice of p there exist integers ξ^c_{ik} with $|\xi^c_{ik}| \le p$ so that

$$\sum_{i\in\bigcup_{j\le k} I_j} \xi^c_{ik}\boldsymbol{a}^i + c\sum_{i\in I_{k+1}} \boldsymbol{a}^i = \boldsymbol{0}.$$

Multiplying this equation with x_{k+1} and adding it to the first one yields

$$\sum_{i\in\bigcup_{j\le k} I_j} (\xi^c_{ik}x_{k+1} + y_i)\boldsymbol{a}^i + \sum_{i\in I_{k+1}} cx_{k+1}\boldsymbol{a}^i = \boldsymbol{0}.$$

Recall that $y_i \in M_{p,c}(x_0, \ldots, x_k)$ for every i and $|\xi_{ik}^c| \leq p$. Hence, $y_i + \xi_{ik}^c x_{k+1} \in M_{p,c}(x_0, \ldots, x_{k+1})$. Obviously, $cx_{k+1} \in M_{p,c}(x_0, \ldots, x_{k+1})$. Thus we have constructed a solution of $A_{k+1}\boldsymbol{x} = \boldsymbol{0}$ which is contained in $M_{p,c}(x_0, \ldots, x_{k+1})$. □

The following partition theorem for (m, p, c)-sets is from Deuber (1973):

Theorem 2.11 (Deuber). *Let m, p, c and r be positive integers. Then there exist positive integers n, q and d such that for every coloring $\Delta : \mathbb{N} \to r$ of the positive integers every (n, q, d)-set $N \subseteq \mathbb{N}$ contains a monochromatic (m, p, c)-set.*

Combining this result with Lemma 2.10 proves the partition theoretic part of Rado's theorem, viz. that A having the column property implies that $A\boldsymbol{x} = \boldsymbol{0}$ is partition regular. In fact Theorem 2.11 is stronger than needed for our purposes. Deuber used this full partition theorem for (m, p, c)-sets to answer a conjecture of Rado:

A subset $S \subseteq \mathbb{N}$ is called partition regular if every partition regular system of equations is solvable in S. Deuber showed that if S is partition regular and S is colored with finitely many colors then one of the color classes is again partition regular.

Originally, Theorem 2.11 was proved with the help of van der Waerden's theorem on arithmetic progressions. Later, Leeb (1975) observed that the use of Hales-Jewett's theorem provides a more elegant proof.

A proof of Deuber's theorem based on Hales-Jewett's theorem will be given in Sect. 4.2.

2.5.3 *Proof of Rado's Theorem*

Deuber's Theorem 2.11 together with Lemma 2.10 implies that the column property of A implies that $A\boldsymbol{x} = \boldsymbol{0}$ is partition regular. Hence, it remains to show that the partition regularity of $A\boldsymbol{x} = \boldsymbol{0}$ implies the column property of A.

Let $A = (\boldsymbol{a}^0, \ldots, \boldsymbol{a}^{n-1})$ be a $k \times n$-matrix such that $A\boldsymbol{x} = \boldsymbol{0}$ is partition regular.

Let $I \subseteq n$ and $\boldsymbol{a} \neq 0$ be a vector in $\mathbb{Z}^k$ that is not a (rational) linear combination of the $\boldsymbol{a}^i, i \in I$. Let $P(I, \boldsymbol{a})$ be the set of all primes p such that for some nonnegative integer m we have that $p^m \cdot \boldsymbol{a}$ is a linear combination of the $\boldsymbol{a}^i, i \in I$, modulo p^{m+1}. Then $P(I, \boldsymbol{a})$ is finite.

To see this let $\boldsymbol{b} \in \mathbb{Q}^k$ be such that $\boldsymbol{b}^T \cdot \boldsymbol{a}^i = 0$ for every $i \in I$ but $\boldsymbol{b}^T \cdot \boldsymbol{a} \neq 0$. Without loss of generality we can assume that $\boldsymbol{b} \in \mathbb{Z}^k$ and, hence, $\boldsymbol{b}^T \cdot \boldsymbol{a} \in \mathbb{Z}$.

Let m be some nonnegative integer. Then

$$p^m \boldsymbol{a} = \sum_{i \in I} \xi_i \boldsymbol{a}^i \pmod{p^{m+1}}$$

implies that

$$\boldsymbol{b}^T p^m \boldsymbol{a} = 0 \pmod{p^{m+1}}.$$

Hence, $p \mid b^T \cdot \boldsymbol{a}$ which is only true for finitely many primes.

Now choose a prime p which is not in $P(I, \boldsymbol{a})$ for every $\boldsymbol{a} = \sum_{j \in J} \boldsymbol{a}^j$ where $J \subseteq n$ and $\boldsymbol{a}$ is not a linear combination of $\boldsymbol{a}^i, i \in I$. Moreover, let p be not one of the finitely many primes which have the property that $\sum_{i \in I} \boldsymbol{a}^i \equiv \boldsymbol{0} \pmod{p}$ for some $I \subseteq n$ with $\sum_{i \in I} \boldsymbol{a}^i \neq 0$.

Every positive integer x admits a unique representation as $x = y(x) p^{z(x)}$ where $y(x) \not\equiv 0 \pmod{p}$. Let $\Delta_p : \mathbb{N} \to [1, p-1]$ be the coloring given by $\Delta_{\mathrm{p}}(x) = y(x) \pmod{p}$. Since $A\boldsymbol{x} = \boldsymbol{0}$ is partition regular there exists a solution which is monochromatic with respect to Δ_p. This solution has the form

$$x_i = p^{z(x_i)}(p\alpha(x_i) + r) \quad \text{for every } i < n,$$

where $r \in [1, p-1]$ is the same for every i. Without loss of generality we can assume that

$$z(x_0) \leq \ldots \leq z(x_{n-1}).$$

We will partition n according to the $z(x_i)$-values and show that this partition proves that A has the column property. For this purpose let

$$\begin{aligned}
m_0 &= z(x_0) = \ldots = z(x_{i_1}) \\
m_1 &= z(x_{i_1+1}) = \ldots = z(x_{i_2}) \\
&\vdots \\
m_\ell &= z(x_\ell + 1) = \ldots = z(x_{n-1}) \text{ and} \\
m_0 &< m_1 < \ldots < m_\ell
\end{aligned}$$

Now put

$$\begin{aligned}
I_0 &= \{0, \ldots, i_1\} \\
I_1 &= \{i_1 + 1, \ldots, i_2\} \\
&\vdots \\
I_\ell &= \{i_t + 1, \ldots, n-1\}.
\end{aligned}$$

First we verify that $\sum_{i\in I_0} \boldsymbol{a}^i = 0$. Since $x_0, \ldots, x_{n-1}$ is a solution we have that

$$\sum_{i<n} x_i \boldsymbol{a}^i = \mathbf{0}.$$

Thus in particular

$$\sum_{i\in I_0} x_i \boldsymbol{a}^{\mathfrak{i}} + \sum_{i\in n\setminus I_0} x_i \boldsymbol{a}^i \equiv \mathbf{0} \pmod{p^{m_0+1}}.$$

For $i \in n \setminus I_0$ we have that $x_i \equiv 0 \pmod{p^{m_0+1}}$ and for every $i \in I_0$ that $x_i = p^{m_0}(p\alpha(x_i) + r)$. Hence

$$r \cdot \sum_{i\in I_0} \boldsymbol{a}^i \equiv \mathbf{0} \pmod{p}.$$

Since $r \in [1, p-1]$ and by choice of p it follows that $\sum_{i\in I_0} \boldsymbol{a}^i = \mathbf{0}$.

Now we verify along the same lines that for $k > 0$ the sum of the columns in class I_k is a linear combination of the columns in the previous classes. As before, we have that

$$\sum_{i\in\bigcup_{j<k} I_j} x_i \boldsymbol{a}^i + \sum_{i\in I_k} x_i \boldsymbol{a}^i + \sum_{\in\bigcup_{k<j} I_j} x_i \boldsymbol{a}^i \equiv \mathbf{0} \pmod{p^{m_k+1}}.$$

Hence, reducing modulo p gives

$$\sum_{i\in\bigcup_{j<k} I_j} x_i \boldsymbol{a}^i + rp^{m_k} \sum_{i\in I_k} \boldsymbol{a}^i \equiv \mathbf{0} \pmod{p^{m_k+1}}.$$

Thus by choice of p we obtain the desired result, completing the proof of Rado's theorem. □

It should be mentioned that Furstenberg (1981) obtained a proof of Rado's theorem of completely different nature using methods from topological dynamics.

2.5.4 Finite and Infinite Sums

Of course, Rado's theorem covers Hilbert's cube lemma, Schur's lemma and van der Waerden's theorem as well as Schur's extension of it. Because of its particular interest we will briefly discuss one other special case of Rado's theorem.

Recalling Example (3) we get as an immediate consequence of Rado's theorem the following finite sum theorem:

Theorem 2.12 (Rado, Folkman, Sanders). *Let m and r be positive integers. Then there exists a least positive integer $n = FS(m,r)$ such that for every coloring $\Delta : [1,n] \to r$ there are m positive integers $a_0, \ldots, a_{m-1}$ such that for all nonempty sets $I, J \subseteq m$ it follows that*

$$\Delta(\sum_{i\in I} a_i) = \Delta(\sum_{j\in J} a_j).$$

Theorem 2.12 was rediscovered several times, among others by Folkman (see Graham 1981 or Graham et al. 1980) and Sanders (1968) leading to the present name of this theorem.

Folkman's proof uses van der Waerden's theorem. The idea of the second proof of Schur's lemma (cf. Sect. 2.2) has been extended by Nešetřil and Rödl (1983a) to obtain a proof of the Rado-Folkman-Sanders theorem from Ramsey's theorem. In Sect. 5.2.4 we will get the finite sum theorem as an immediate application of Hales-Jewett's theorem.

Another combinatorial proof of the finite sum theorem was given by Taylor (1981). His proof is remarkable because it provides the least known upper bound on $FS(m,r)$, viz.

$$\mathrm{FS} \leq 2^{r^{3^{\cdot^{\cdot^{\cdot^{r^3}}}}}} \Big\}^{2r(m-1)}, \quad m, r \geq 2.$$

Having the finite sum theorem in hands it is natural to ask whether or not an infinite version of it is valid. Graham and Rothschild (1971) conjectured an infinite generalization of the Rado-Folkman-Sanders theorem which was proved by Hindman (1974):

Theorem 2.13 (Hindman). *Let r be a positive integer. Then for every coloring $\Delta : \mathbb{N} \to r$ there exist infinitely many integers $a_0, a_1, a_2, \ldots$ such that for all nonempty finite sets $I, J \subseteq \omega$ it follows that*

$$\Delta(\sum_{i\in I} a_i) = \Delta(\sum_{j\in J} a_j),$$

i.e., all finite sums of the a_i get the same color.

Several proofs have been given for this theorem, e.g., by Baumgartner (1974) using some kind of combinatorial forcing, by Glazer (see, Hindman 1979) using idempotent ultrafilters in $\beta\mathbb{N}$ and by Furstenberg and Weiss (1978) using topological dynamics. The reader may consult one of these references for a proof of Hindman's theorem.

Assuming the axiom of choice it is easy to see that (coloring the reals) one cannot expect to get also infinite sums in the same color. Of course, taking infinite sums necessarily requires convergence. But restricting to in a sense constructive colorings,

viz. colorings having the property that each color class has the property of Baire, it can be proved that there always exists an infinite sequence of reals (whose sum converges) such that all their sum (finite or infinite, but without repetition) get the same color (Prömel and Voigt 1990).

Springer

http://www.springer.com/978-3-319-01314-5

Ramsey Theory for Discrete Structures
Prömel, H.J.
2013, XVI, 232 p. 13 illus., Hardcover
ISBN: 978-3-319-01314-5