Teil (i). Jede Box lässt sich als Paar (b, r) darstellen, wobei b die Anzahl der blauen und r die Anzahl der roten Bälle bezeichnet. Da beim zufälligen Ziehen eines Balls jeder der $b + r$ Bälle gleichwahrscheinlich ist, ergibt sich

$$P = \frac{b}{b+r} \cdot \frac{b-1}{b-1+r} = \frac{b(b-1)}{(b+r)(b+r-1)}$$

als Wahrscheinlichkeit für zwei blaue Bälle bei zufälliger Entnahme zweier Bälle. Zur Vereinfachung sei $n = b + r$ die Gesamtzahl der Bälle. Eine Box (b, r) ist nach Aufgabenstellung magisch, wenn $P = 50\,\%$ gilt, d. h. wenn

$$P = \frac{b(b-1)}{(b+r)(b+r-1)} = \frac{b(b-1)}{n(n-1)} = \frac{1}{2}$$

erfüllt ist. Umformen liefert folgende quadratische Gleichung in b:

$$\frac{b(b-1)}{n(n-1)} = \frac{1}{2} \Leftrightarrow b^2 - b - \frac{1}{2}n(n-1) = 0.$$

Mit einer Lösungsformel oder quadratischer Ergänzung ergeben sich

$$b_{1/2} = \frac{1}{2} \pm \sqrt{\frac{1}{4} + \frac{1}{2}n(n-1)} = \frac{1}{2}\left(1 \pm \sqrt{1 + 2n(n-1)}\right)$$

als Lösungen. Abbildung 1 zeigt einen Plot der daraus entstehenden Funktion. Jeder magischen Box kann ein Gitterpunkt[1] mit $n \geq 2$ zugeordnet werden und umgekehrt. Dabei brauchen wir b_2 nicht zu betrachten, denn für $n \geq 2$ ist

$$b_2 \leq \frac{1}{2}\left(1 - \sqrt{1 + 2 \cdot 2 \cdot (2-1)}\right) = \frac{1-\sqrt{5}}{2} < 0.$$

Sei $R = 1 + 2n(n-1)$ der Radikand und $N = 10^6$ die obere Grenze, bis zu der wir magische Boxen suchen. Dann genügt es, für $n \leq N$ die Gleichung

$$b = \frac{1}{2}\left(1 + \sqrt{R}\right)$$

zu untersuchen. Offenbar ist b genau dann eine positive ganze Zahl, wenn R eine ungerade Quadratzahl ist. Denn ist R eine gerade Quadratzahl, dann ist $\sqrt{R}$ gerade, $1 + \sqrt{R}$ ungerade und damit b keine ganze Zahl. Ist R gar keine Quadratzahl, dann ist $\sqrt{R}$ und damit auch b irrational, also keine ganze Zahl.

An dieser Stelle ist ein Brute-Force-Ansatz zielführend: Für $2 \leq n \leq N$ prüfen wir, ob R eine ungerade Quadratzahl ist und erhalten ggf. $(b, n - b)$ als Lösung.

Sei R gegeben. Wie überprüfen wir, ob R eine ungerade Quadratzahl ist?

[1] Punkt mit ganzzahligen Koordinaten bzw. Element der Menge $\mathbb{Z} \times \mathbb{Z}$

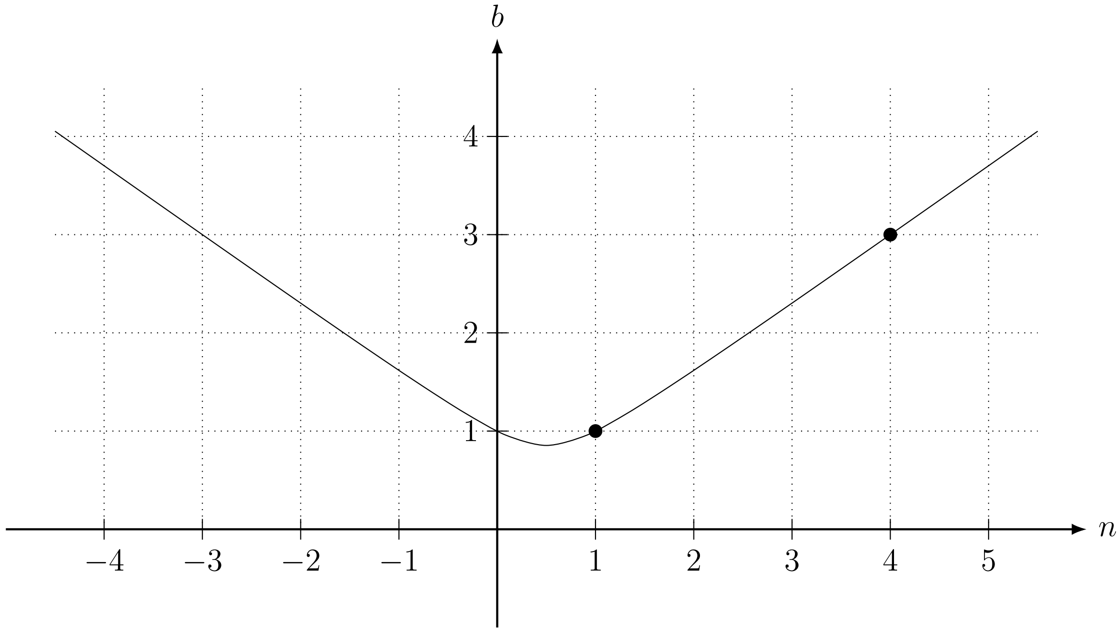

Abbildung 1: Lösungen sind Gitterpunkte auf dem Funktionsgraphen

Erster Ansatz. Nach dem Prinzip der linearen Suche überprüfen wir für alle k mit $1 \leq k < R$, ob $R = k^2$ gilt. Insgesamt sind dann

$$\sum_{n=2}^{N}(R-1) = \sum_{n=2}^{N} 2n(n-1) = \frac{2}{3}N(N-1)(N+1) \in \mathcal{O}(N^3)$$

solche Tests erforderlich. Im Erfolgsfall haben wir $\sqrt{R}$ direkt bestimmt, was aber auch die folgenden drei Ansätze leisten. Für $N \leq 2\,000$ ist die Lösung passabel, größere Eingaben bringen sie aber wegen der kubischen Laufzeit schnell an ihre Grenzen. Die Lösung ist trotzdem ein guter Start für Optimierungen.

Zweiter Ansatz. Die Quadratwurzelfunktion ist streng monoton steigend, d. h. für $x > y$ ist stets $\sqrt{x} > \sqrt{y}$. Somit genügt es, wie im ersten Ansatz vorzugehen und abzubrechen, sobald $k^2 > R$ gilt. Dann ist nämlich auch $(k+1)^2 > R$, $(k+2)^2 > R$, $(k+3)^2 > R$, $\ldots$. Diese Optimierung reduziert die Anzahl der Tests auf $1 + \sqrt{R}$ je n, sodass insgesamt nur noch

$$\sum_{n=2}^{N}\left(1+\sqrt{1+2n(n-1)}\right) \leq \sum_{n=1}^{N}\sqrt{4n^2} = \sum_{n=1}^{N} 2n = N(N+1) \in \mathcal{O}(N^2)$$

Tests notwendig sind und die Laufzeit um den Faktor N reduziert ist. Ist $N = 10^6$, dann läuft das Programm ca. 1 Million mal schneller. Der Ansatz ist für $N \leq 100\,000$ passabel und damit für eine vollständige Lösung der Teilaufgabe zu langsam.

Dritter Ansatz. Mit der binären Suche bzw. dem Intervallhalbierungsverfahren ist eine Lösung mit Laufzeit $\mathcal{O}(N \log N)$ möglich, die für $N = 10^6$ ausreicht.

Dazu starten wir mit dem Intervall $[1; R]$, berechnen den abgerundeten Mittelwert der Intervallgrenzen $m = \left\lfloor \frac{1+R}{2} \right\rfloor$ und unterscheiden drei Fälle:

- $m^2 = R$: Wir haben $m = \sqrt{R}$ gefunden und sind fertig.
- $m^2 > R$: Dann ist $\sqrt{R} < m$ und wir setzen die Suche im Intervall $[1; m-1]$ auf analoge Weise fort.
- $m^2 < R$: Dann ist $\sqrt{R} > m$ und wir setzen die Suche im Intervall $[m+1; R]$ auf analoge Weise fort.

In jedem Schritt halbieren wir die Größe des zu untersuchenden Intervalls, sodass nach k Schritten nur noch $1/2^k$ aller möglichen Kandidaten für $\sqrt{R}$ zu untersuchen sind. Da es am Anfang R Kandidaten gibt, ist nach spätestens $1 + \log_2 R$ Schritten Schluss, dann sind nämlich nur noch

$$R \cdot \frac{1}{2^{1+\log_2 R}} = R \cdot \frac{1}{2} \cdot \frac{1}{2^{\log_2 R}} = \frac{R}{2} \cdot \frac{1}{R} = \frac{1}{2} < 1,$$

Kandidaten, also keine Kandidaten mehr übrig. Da für jedes n mit $2 \leq n \leq N$ stets $R \leq 2N^2$ gilt, sind wegen $\log_2(2N^2) = 1 + 2\log_2 N$ nur jeweils $\mathcal{O}(\log N)$ Schritte und insgesamt $\mathcal{O}(N \log N)$ Schritte erforderlich. Für $N \leq 3 \cdot 10^8$ genügt dies.

Implementierung in C++:

```cpp
typedef unsigned long long num;
int main() {
    for (num n = 2; n <= 1000 * 1000; ++n) {
        num r = 1 + 2 * n * (n - 1);
        if (r % 2 == 0) continue;
        num low = 1, high = r, rs = -1;
        while (low <= high && rs == -1) {
            num m = (low + high) / 2;
            num ms = m * m;
            if (ms == r) rs = m;
            if (ms < r) low = m + 1;
            if (ms > r) high = m - 1;
        }

        if (rs != -1) {
            num b = (1 + rs) / 2;
            cout << "(" << b << ", " << (n - b) << ")" << endl;
        }
    }

    return 0;
}
```

Hinweis. Der Datentyp `unsigned long long` dient der Vermeidung von Überlaufen.

Es gibt nur acht magische Boxen mit höchstens 1 000 000 Bällen und diese sind in Tabelle 1 dargestellt. Obwohl die Zahlenwerte keinem einfachen Muster folgen, ist zu erkennen, dass diese sehr schnell (genauer: exponentiell) wachsen. Der kritische Leser mag verifizieren, dass die angegebenen Boxen tatsächlich magisch sind.

Nr.	Blaue Bälle	Rote Bälle	Summe
1	3	1	4
2	15	6	21
3	85	35	120
4	493	204	697
5	2 871	1 189	4 060
6	16 731	6 930	23 661
7	97 513	40 391	137 904
8	568 345	235 116	803 761

Tabelle 1: Antwort zu Teil (i) – Magische Boxen mit $n \leq 10^6$

Vierter Ansatz. Überraschenderweise gibt es sogar eine sehr kurze $\mathcal{O}(N)$-Lösung, die aber nicht ganz einfach zu finden ist. Statt für jedes n unabhängig zu überprüfen, ob R eine ungerade Quadratzahl ist, können wir ausnutzen, dass R mit n streng monoton steigt, d. h. für $n' > n$ ist $R' > R$. Somit genügt es, eine Variable k mit dem Wert $k = 1$ zu initialisieren, sodass die Invarianten $k^2 \leq R$ und k ungerade erfüllt sind. Beim Iterieren über alle $n \leq N$ inkrementieren wir k zunächst solange um 2, bis $k^2 < R$ verletzt ist. Gilt $k^2 = R$, dann ist R eine ungerade Quadratzahl mit $k = \sqrt{R}$. Gilt $k^2 > R$, dann ist R keine ungerade Quadratzahl.

Implementierung in C++:

```
typedef unsigned long long num;
int main() {
    num k = 1;
    for (num n = 2; n <= 1000 * 1000; ++n) {
        num r = 1 + 2 * n * (n - 1);
        while (k * k < r) k += 2;
        if (k * k == r) {
            num b = (1 + k) / 2;
            cout << "(" << b << ", " << (n - b) << ")" << endl;
        }
    }

    return 0;
}
```

Teil (ii). Zauberhafte Boxen existieren nicht. Wir führen einen Widerspruchsbeweis und nehmen an, es gäbe eine zauberhafte Box mit b blauen Bällen und n Bällen insgesamt. Dann ist die Wahrscheinlichkeit für zwei blaue Bälle analog zu Teil (i)

$$P = \frac{b(b-1)}{n(n-1)} = \frac{1}{4},$$

was sich äquivalent umformen lässt zur quadratischen Gleichung

$$b^2 - b - \frac{1}{4}n(n-1) = 0.$$

Mit einer Lösungsformel oder quadratischer Ergänzung ergeben sich

$$b_{1/2} = \frac{1}{2} \pm \sqrt{\frac{1}{4} + \frac{1}{4}n(n-1)} = \frac{1}{2}\left(1 \pm \sqrt{1+n(n-1)}\right).$$

als Lösungen in b. Da b nach Annahme eine ganze Zahl ist, sind auch diese Lösungen ganzzahlig und folglich ist der Radikand $1+n(n-1)$ eine Quadratzahl. Nun gelten aber für $n \geq 2$ die Ungleichungen

$$(n-1)^2 < 1 + n(n-1) < n^2,$$

denn es ist $(n-1)^2 = n^2 - 2n - 1 < n^2 - n + 1 = 1 + n(n-1) < n^2$. Somit liegt $1 + n(n-1)$ zwischen zwei aufeinanderfolgenden Quadratzahlen und kann daher nicht selbst eine Quadratzahl sein. Im Widerspruch zur Annahme ist b keine ganze Zahl und es folgt, dass die Annahme falsch ist. Es gibt also keine zauberhafte Box.

Teil (iii). Mit den Überlegungen aus Teil (i) genügt es zu zeigen, dass die Gleichung

$$b^2 - b - \frac{1}{2}n(n-1) = 0$$

unendlich viele Lösungen in positiven ganzen Zahlen n und b hat. Durch scharfes Hinsehen und Probieren finden wir, dass

$$\left(b-\frac{1}{2}\right)^2 - \frac{1}{2}\left(n-\frac{1}{2}\right)^2 = \underbrace{b^2 - b - \frac{1}{2}n(n-1)}_{=0} + \frac{1}{4} - \frac{1}{8} = \frac{1}{8}$$

gilt. Multiplikation mit 8 liefert die äquivalente Gleichung

$$2\,(2b-1)^2 - (2n-1)^2 = 1.$$

Mit der Substitution $x \leftarrow 2n-1$ und $y \leftarrow 2b-1$ ergibt sich die Gleichung

$$2y^2 - x^2 = 1 \Leftrightarrow x^2 - 2y^2 = -1.$$

Gesucht sind Lösungen in ungeraden positiven ganzen Zahlen x und y, denn es sind

$$n = \frac{x+1}{2} \text{ und } b = \frac{y+1}{2}$$

genau in diesen Fällen positive ganze Zahlen. Die Lösung $(x, y) = (1, 1)$ beschreibt wegen $b = 1 < 2$ keine magische Box. Trotzdem kann diese sogenannte *Basislösung* zur Konstruktion unendlich vieler weiterer Lösungen verwendet werden.

Sei (x, y) eine beliebige Lösung. Dann folgt der Reihe nach

$$\begin{aligned}(-1)^2 \cdot (x^2 - 2y^2) &= -1 \\ \Leftrightarrow (1^2 - 2 \cdot 1^2)^2 \cdot (x^2 - 2y^2) &= -1 \\ \Leftrightarrow \Big((1 - \sqrt{2})(1 + \sqrt{2})\Big)^2 \cdot (x - \sqrt{2}y)(x + \sqrt{2}y) &= -1 \\ \Leftrightarrow (x - \sqrt{2}y)(1 - \sqrt{2})^2 \cdot (x + \sqrt{2}y)(1 + \sqrt{2})^2 &= -1 \\ \Leftrightarrow (3 - 2\sqrt{2})(x - \sqrt{2}y) \cdot (3 + 2\sqrt{2})(x + \sqrt{2}y) &= -1 \\ \Leftrightarrow \Big(3x + 4y - \sqrt{2}(2x + 3y)\Big)\Big(3x + 4y + \sqrt{2}\,(2x + 3y)\Big) &= -1 \\ \Leftrightarrow (3x + 4y)^2 - 2(2x + 3y)^2 &= -1,\end{aligned}$$

sodass mit (x, y) auch $(x', y') = (3x + 4y, 2x + 3y)$ eine Lösung ist. Tatsächlich ist für x und y ungerade sowohl $3x + 4y$ als auch $2x + 3y$ ungerade. Mit $x = 2n - 1$ und $y = 2b - 1$ ergeben sich die Lösungen:

$$\begin{aligned}x' &= 3x + 4y = 3(2n - 1) + 4(2b - 1) = 6n + 8b - 7 \\ y' &= 2x + 3y = 2(2n - 1) + 3(2b - 1) = 4n + 6b - 5\end{aligned}$$

Durch die Rücksubstitution folgen somit die Rekursionsgleichungen

$$\begin{aligned}n' &= \frac{6n + 8b - 7 + 1}{2} = 3n + 4b - 3, \\ b' &= \frac{4n + 6b - 5 + 1}{2} = 2n + 3b - 2.\end{aligned}$$

Starten wir mit $(n, b) = (4, 3)$, dann lassen sich unendlich viele Lösungen mithilfe dieser Rekursionsgleichungen berechnen. Damit ist der Beweis abgeschlossen.

Anmerkung 1. Zur Probe kann mit vollständiger Induktion bewiesen werden, dass alle durch die Rekursionsgleichungen erzeugten Boxen tatsächlich sämtlich magisch und voneinander verschieden (z. B. in n streng monoton steigende Folge) sind. Der Beweis ist jedoch trivial bzw. erfordert lediglich Rechenarbeit.

Anmerkung 2. Geschicktes Einsetzen und Untersuchen charakteristischer Polynome oder Eigenwertdekomposition liefern explizite Formeln für n und b der k-ten Box:

$$\begin{aligned}n_k &= \frac{1}{2} + \frac{(1 + \sqrt{2})^{2k-1} + (1 - \sqrt{2})^{2k-1}}{4} \\ b_k &= \frac{1}{2} + \frac{(1 + \sqrt{2})^{2k-1} - (1 - \sqrt{2})^{2k-1}}{4\sqrt{2}}\end{aligned}$$

Anmerkung 3. Die Rekursionsgleichungen und expliziten Formeln aus Anmerkung 2 beschreiben nur eine unendliche Teilmenge der magischen Boxen, sondern sogar alle magischen Boxen! Der Beweis sei dem Leser als Übung überlassen.