

**Zeitschrift:** Mitteilungen / Vereinigung Schweizerischer Versicherungsmathematiker  
= Bulletin / Association des Actuaires Suisses = Bulletin / Association of  
Swiss Actuaries

**Herausgeber:** Vereinigung Schweizerischer Versicherungsmathematiker

**Band:** 61 (1961)

**Artikel:** Die Zahl der Vergleichsoperationen beim Mischen auf 3 Bahnen

**Autor:** Boehm, Carl

**DOI:** <https://doi.org/10.5169/seals-966729>

### **Nutzungsbedingungen**

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

### **Conditions d'utilisation**

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

### **Terms of use**

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

**Download PDF:** 15.03.2025

**ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>**

## Die Zahl der Vergleichsoperationen beim Mischen auf 3 Bahnen

*Von Carl Boehm, Düsseldorf*

### Zusammenfassung

Der Verfasser bestimmt den Erwartungswert für die Zahl der Vergleichsoperationen beim *Mischen auf 3 Bahnen* von  $N = 3n$  Sortiereinheiten mit

$$\varphi^{(3)} = \frac{2N^2}{N+3} \frac{5N+9}{6N+9}$$

und leitet daraus ab als Erwartungswert der Zahl der Vergleichsoperationen für das *Sortieren durch fortgesetztes Mischen auf 3 Bahnen* (straight merging sort on 3 ways) bei  $N = 3^k$  ( $k \geq 5$ )

$$\Phi^{(3)} = 2N (3,475873 + (\log_3 N - 5) R)$$

mit

$$0,83 < R < 0,8\bar{3}.$$

Dies ergibt eine *grössere* Zahl von Vergleichen je Einheit als beim Sortierprozess auf 2 Bahnen.

Überraschend ist, dass das Mischen auf 2 Bahnen von zuerst  $n$  und  $n$  Einheiten und anschliessend von  $2n$  und  $n$  Einheiten zum *gleichen* Erwartungswert wie das Mischen auf 3 Bahnen führt. Dies ermöglicht die Angabe brauchbarer oberer Grenzen für die Zahl der Vergleichsoperationen beim Mischen auf mehr als 3 Bahnen.

Das Mischen von 2 Sequenzen mit je  $n$  «Sortiereinheiten» erfordert [1], [2]<sup>1)</sup> im Durchschnitt

$$\varphi^{(2)} = \frac{2n^2}{n+1} = \frac{N^2}{N+2} \quad (N = 2n). \quad (1)$$

Vergleiche von je 2 «Sortiermerkmalen». Daraus folgt, dass das *Sortieren durch fortgesetztes Mischen mit bekannter Sequenzlänge auf 2 Bahnen* (straight merging sort on 2 ways) von  $N = 2^k$  Sortiereinheiten durchschnittlich

---

<sup>1)</sup> Die in eckigen Klammern angegebenen Zahlen beziehen sich auf das Literaturverzeichnis.

$$\begin{aligned} \Phi^{(2)} &= 2^k k - 2^k \sum_{i=1}^k \frac{1}{2^{i-1} + 1} \\ &= N \log_2 N - N \sum_{i=1}^{\log_2 N} \frac{1}{2^{i-1} + 1} \end{aligned} \quad (2)$$

Vergleichsoperationen benötigt.

Im folgenden werden die entsprechenden Erwartungswerte für das *Mischen auf 3 Bahnen* genau bestimmt und gezeigt, dass sie, wenn auch nur wenig, grösser sind. Schliesslich werden noch obere Schranken dieser Werte bei Verwendung von mehr als 3 Bahnen angegeben.

1. Gegeben seien 3 Folgen von je  $n$  Sortiereinheiten ( $N = 3n$ ), innerhalb jeder Folge seien die Sortiereinheiten nach nicht-abnehmenden Sortiermerkmalen geordnet (Sequenzen). Durch *einen* Mischvorgang [3], [4] wird hieraus *eine* Sequenz der «Länge»  $N$  erzeugt. Wieviel Vergleiche von je 2 Sortiermerkmalen sind hierfür im Durchschnitt aller möglichen Permutationen erforderlich?

Die Gesamtzahl aller Permutationen beträgt unter Berücksichtigung der Sequenzeigenschaft der 3 Folgen

$$(3n)! / (n!)^3 \quad (3)$$

in bezug auf diese Gesamtheit ist unter der Annahme der Gleichwahrscheinlichkeit jeder dieser Permutationen der Erwartungswert der Zahl der Vergleichsoperationen zu bestimmen.

2. Für die Berechnung nehmen wir *zunächst* an, dass die Bestimmung der Sortiereinheit mit kleinstem Merkmal unter 3 Sortiereinheiten immer 2 Vergleiche benötigt; bei 2 Sortiereinheiten ist immer 1 Vergleich erforderlich.

Unter dieser Annahme lässt sich leicht eine *untere* und *obere* Grenze des gesuchten Erwartungswertes angeben:

a) Die *untere* Grenze wird erreicht, wenn die Sequenzen sich nicht «überlagern», wenn also gilt:

(alle Sortiermerkmale der 1. Sequenz) < (alle Sortiermerkmale der 2. Sequenz) < (alle Sortiermerkmale der 3. Sequenz).

Dann ist die Zahl der Vergleichsoperationen

$$= 2n + n = 3n = N.$$

Dieser Fall hat die Häufigkeit  $3! (n!)^3 / (3n)!$ .



Jede der möglichen Kombinationen von  $k, i, l$ -Werten kommt

$$2 \cdot 3 \cdot \frac{(3n-k-1)! (i-1)!}{(n-1)! (n-l)! (n-(k-l))! (l-1)! (i-(k-l))!} \text{-mal vor,}$$

oder, da  $(n-l) + (n-(k-l)) = 2n-k$   
 und  $(2n-k) + (n-1) = 3n-k-1$

$$= 2 \cdot 3 \cdot \binom{3n-k-1}{n-1} \binom{2n-k}{n-l} \binom{i-1}{l-1}. \quad (4)$$

Der Faktor 3 erklärt sich aus den 3 Möglichkeiten für das Zeichen, dessen Vorrat zuerst erschöpft ist; der Faktor 2 aus der Symmetrie der beiden übrigen.

4. Dass mit diesem Ansatz alle Möglichkeiten erfasst werden, ergibt sich aus

$$\sum_{k=2}^{2n} \left( \sum_{i=\max(1, k-n)}^{k-1} \left( \sum_{l=\max(1, k-n)}^{\min(i, n)} 3! \binom{3n-k-1}{n-1} \binom{2n-k}{n-l} \binom{i-1}{l-1} \right) \right) = \frac{(3n)!}{(n!)^3}.$$

Da 
$$\sum_l \binom{2n-k}{n-l} \binom{i-1}{l-1} = \binom{2n-k+i-1}{n-1}$$

kann man obige Summe wie folgt zerlegen:

$$3! \left\{ \sum_{k=2}^{n+1} \binom{3n-k-1}{n-1} \sum_{i=1}^{k-1} \binom{2n-k+i-1}{n-1} + \sum_{k=n+2}^{2n} \binom{3n-k-1}{n-1} \sum_{i=k-n}^{k-1} \binom{2n-k+i-1}{n-1} \right\}.$$

Hieraus ergibt sich

$$= 3! \left\{ \binom{2n-1}{n} \sum_{k=1}^{2n} \binom{3n-k-1}{n-1} - \sum_{k=2}^{n+1} \binom{3n-k-1}{n-1} \binom{2n-k}{n} \right\}$$

$$= 3! \binom{2n-1}{n} \left\{ \binom{3n-2}{n} - \binom{3n-2}{n-2} \right\} = \frac{(3n)!}{(n!)^3} \text{ w. z. b. w.}$$

5. Jede der möglichen  $(k, i, l)$ -Kombinationen erhalte nun das Gewicht  $(3n-k) \cdot 2 + i \cdot 1 = 6n - 2k + i$ .

Dieser Ansatz erklärt sich aus folgender Überlegung:

Zur Identifikation einer bestimmten Folge, d.h. zur Feststellung, um welche der möglichen Permutationen es sich handelt, müssen durch sukzessive Vergleiche, beginnend links, die einzelnen Zeichen festgestellt werden. Zunächst (A-Bereich) werden nach Annahme 2 Vergleiche gemacht, bis der Vorrat eines Zeichens erschöpft ist. Dann

(B-Bereich) genügt für jede der  $i$ -Stellen 1 Vergleich; im C-Bereich sind keine Vergleiche mehr nötig. Das Gewicht gibt also die Zahl der Vergleichsoperationen an. Der Erwartungswert beträgt dann

$$3! \sum_k \left( \sum_i \left( \sum_l (6n-2k+i) \binom{3n-k-1}{n-1} \binom{2n-k}{n-l} \binom{i-1}{l-1} \right) \right) / \frac{(3n)!}{(n!)^3}.$$

Für die Berechnung geht man analog wie in 4 vor; nach Summation über  $l$  ergibt sich

$$3!n - 2 \cdot 3! \sum_k k \binom{3n-k-1}{n-1} \sum_i \binom{2n-k+i-1}{n-1} / \frac{(3n)!}{(n!)^3} \\ + 3! \sum_k \binom{3n-k-1}{n-1} \sum_i i \binom{2n-k+i-1}{n-1} / \frac{(3n)!}{(n!)^3}.$$

Die beiden ersten Ausdrücke führen nach der Summation auf

$$3!n \left( 1 - 2 \frac{3n+1}{(n+1)(2n+1)} \right), \quad (5)$$

der dritte Ausdruck auf

$$3! \frac{n^2}{(n+1)(2n+1)}. \quad (6)$$

Zusammen ergeben sie das einfache, formal (1) entsprechende Ergebnis

$$\varphi^{(3)} = \frac{3!n^2}{n+1} = \frac{2N^2}{N+3} \quad (7)$$

Dies ist der gesuchte Erwartungswert für die Zahl der Vergleiche beim Mischen von 3 Sequenzen der Länge  $n$ : das obige Modell und die Gewichte entsprechen genau dem Mischvorgang.

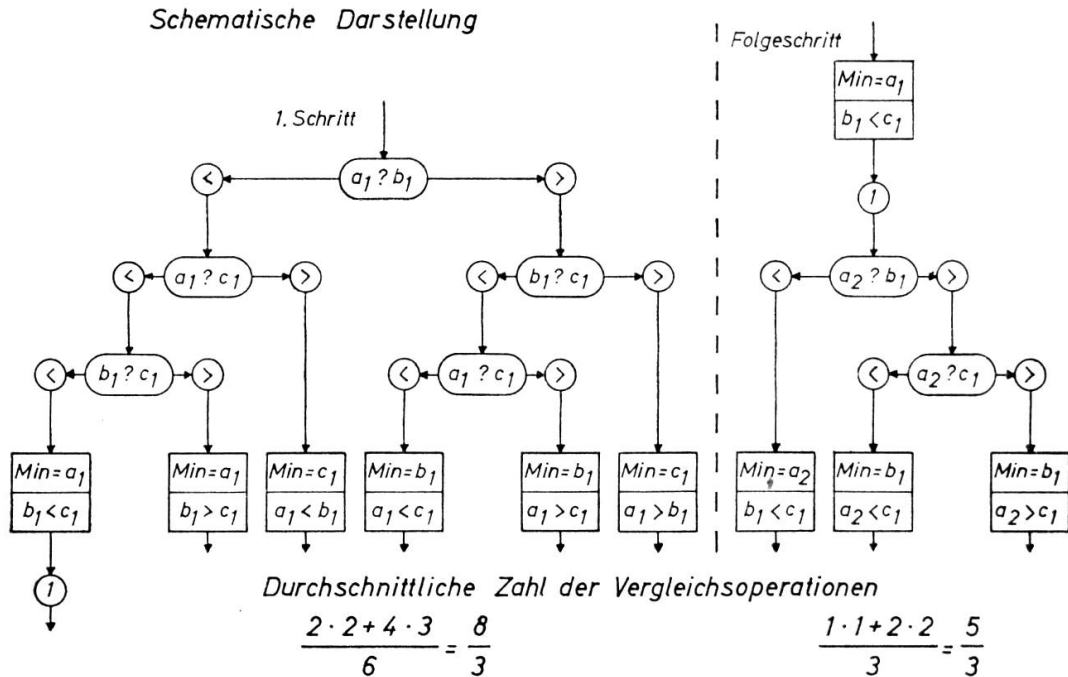
6. Wendet man dieses Ergebnis auf den *Sortierprozess* durch fortgesetztes Mischen auf 3 Bahnen an, so sind für  $N = 3^k$  im  $i$ -ten Durchlauf ( $i = 1, 2, \dots, k$ ) jeweils im Durchschnitt

$$3^{k-i} \frac{2 \cdot 3^{2i}}{3^i + 3}.$$

Vergleichsoperationen nötig, insgesamt also – analog zu (2) –

$$\Phi^{(3)} = 2 \cdot 3^k k - 2 \cdot 3^k \sum_{i=1}^k \frac{1}{3^{i-1} + 1} \quad (8) \\ = 2N \log_3 N - 2N \sum_{i=1}^{\log_3 N} \frac{1}{3^{i-1} + 1}.$$

7. Es wurde bisher angenommen, dass 2 Vergleichsoperationen nötig sind, um die Sortiereinheit mit kleinstem Sortiermerkmal aus 3 Einheiten zu bestimmen. Diese Annahme ist insofern richtig, als immer 2 Vergleiche genügen. Man kann aber durch eine andere Anordnung im Durchschnitt mit weniger Vergleichen auskommen, wie die schematische Darstellung zeigt.



Zu Beginn sind aus jeder der 3 Sequenzen die ersten Einheiten  $a_1, b_1, c_1$  zu vergleichen. Mit 2 oder 3 Vergleichen wird die mit kleinstem Merkmal *und* die Relation zwischen den beiden übrigen bestimmt. Für diesen 1. Schritt sind im Mittel  $8/3$  Vergleiche nötig. Im *Folgeschritt* kann aber die Information über die Relation der beiden übrigen verwendet werden; es sind dann nur noch 1 oder 2 Vergleiche, im Mittel  $5/3$ , zur Bestimmung der Einheit mit kleinstem Merkmal *und* der Relation zwischen den beiden übrigen nötig. Sobald eine Sequenz erschöpft ist, spart man den im 1. Schritt gemachten, zusätzlichen Vergleich wieder ein. In obiger Schreibweise sind daher statt  $6n - 2k + i$  nur noch

$$(3n - k) \frac{5}{3} + i = 5n - \frac{5}{3}k + i. \tag{9}$$

Vergleiche nötig.

Das ergibt im günstigsten Fall ( $k = 2n, i = n$ )

$$8n/3 \text{ Vergleiche (bisher } 3n);$$

im ungünstigsten Fall ( $k = 2, i = 1$ )

$$5n - 7/3 \text{ Vergleiche (bisher } 6n - 3).$$

Setzt man (9) als Gewichte in 5. an, so ergibt sich

$$\varphi^{(3)} = \frac{2n^2}{n+1} \frac{5n+3}{2n+1} = \varphi'^{(3)} \frac{5n+3}{6n+3} = \varphi'^{(3)} \frac{5N+9}{6N+9} \quad (10)$$

was für grössere  $N$  bis zu rund  $1/6$  kleiner als  $\varphi'^{(3)}$  ist.

8. Analog zu 6. ergibt sich für den Sortierprozess als Erwartungswert für die Zahl der Vergleiche bei  $N = 3^k$

$$\Phi^{(3)} = 2 \cdot 3^k \sum_{i=1}^k \frac{3^i}{3^i + 3} \frac{5 \cdot 3^i + 9}{6 \cdot 3^i + 9}.$$

Nimmt man an, dass  $k \geq 5$  ist, so ergibt sich hierfür nach einfachem Ausrechnen der 5 ersten Summanden die Abschätzung

$$\Phi^{(3)} = 2N(3,475873 + (\log_3 N - 5)R), \quad (11)$$

wobei  $0,83 < R < 0,8\bar{3}$ .

9. Ein Zahlenbeispiel:

	2 Bahnen	3 Bahnen
1.	$N = 2^8 = 256$	$N = 3^5 = 243$
	$\Phi^{(2)} = 1726,3$	$\Phi^{(3)} = 1689,3$
	$\frac{\Phi^{(2)}}{N} = 6,743$	$\frac{\Phi^{(3)}}{N} = 6,952$
2.	$N = 2^{16} = 65536$	$N = 3^{10} = 59049$
	$964176 < \Phi^{(2)} < 966216$	$900600 < \Phi^{(2)} < 902569$
	$14,712 < \frac{\Phi^{(2)}}{N} < 14,743$	$15,252 < \frac{\Phi^{(3)}}{N} < 15,285$



10. Zu einer gewissen Überraschung führt folgende Überlegung:

Verallgemeinert man Formel (1) in dem Sinne, dass man auf 2 Bahnen 2 Sequenzen ungleicher Länge, nämlich von  $n$  und  $m$  Einheiten mischt, so ergibt sich nach analogen Überlegungen wie unter 3. die Formel <sup>1)</sup>

$$\varphi_{n,m}^{(2)} = n + m - \frac{n}{m+1} - \frac{m}{n+1}. \quad (1')$$

Ist  $n = m$ , so ergibt sich (1). Es ist nun interessant, festzustellen, dass bei zweimaligem Mischprozess auf 2 Bahnen, nämlich

- a) von 2 Sequenzen der Länge  $n$  und
- b) von 2 Sequenzen der Länge  $n$  und  $2n$ ,

die durchschnittliche Zahl der Vergleiche *dieselbe* ist wie beim Mischen von 3 Sequenzen der Länge  $n$ . In der Tat ist

$$\begin{aligned} \varphi_{n,n}^{(2)} + \varphi_{n,2n}^{(2)} &= \frac{2n^2}{n+1} + 3n - \frac{n}{2n+1} - \frac{2n}{n+1} \\ &= \frac{2n^2}{n+1} \frac{5n+3}{2n+1} = \varphi_{n,n,n}^{(3)}. \end{aligned}$$

Das bedeutet, dass *in bezug auf die Zahl der Vergleichsoperationen* durch zweimaliges Mischen auf 2 Bahnen derselbe Effekt erzielt wird wie durch einmaliges Mischen auf 3 Bahnen. Dies ist darin begründet, dass nur Alternativvergleiche möglich sind.

11. Diese Beobachtung legt die Vermutung nahe, dass sich durch Kombination von  $\varphi_{n,n}^{(2)}$ ,  $\varphi_{n,m}^{(2)}$  und  $\varphi_{n,n,n}^{(3)}$  brauchbare obere Schranken für den Erwartungswert beim Mischen auf mehr als 3 Bahnen ergeben.

Beim Mischen auf z. B. 4 Bahnen wird man zunächst zweimal je 2 Sequenzen der Länge  $n$  und dann einmal 2 Sequenzen der Länge  $2n$  mischen. Dies ergibt ( $N = 4n$ )

$$\varphi_{n,n,n,n}^{(4)} < 2\varphi_{n,n}^{(2)} + \varphi_{2n,2n}^{(2)} = \frac{4n^2}{n+1} \frac{4n+3}{2n+1} = \frac{2N^2(N+3)}{(N+2)(N+4)}.$$

Verfolgt man diesen Gedanken weiter, so ergeben sich Grenzen, die nur sehr wenig unterhalb der Zahl der Vergleichsoperationen im «ungünstigsten Fall» liegen, d. h. wenn man annimmt, dass bis zum

---

<sup>1)</sup> Diese Formel wurde zuerst von A. Schönhage (Köln) abgeleitet.

Ende in *jeder* Sequenz noch Einheiten sind. Bezeichnet  $\varepsilon_b$  die Mindestzahl der Vergleiche, die gemacht werden müssen, um aus  $b$  Einheiten die mit kleinstem Merkmal zu bestimmen (es ist  $\varepsilon_2 = 1$ ,  $\varepsilon_3 = 5/3$ ), so sind in diesem Fall

$$\varphi^{(b)} < (bn - (b-1)) \varepsilon_b + \sum_{v=2}^{b-1} \varepsilon_v \sim bn \varepsilon_b = N \varepsilon_b.$$

Vergleiche nötig. Für  $b = 4$  ist  $\varepsilon_4 = 2$ , und es wird so z. B. für  $n = 100$  die sich ergebende Zahl der Vergleiche 800, während obiger Ansatz 794 ergeben hätte. Für  $b = 8$  ist  $\varepsilon_8 = 3$ ; für  $n = 100$  lauten die entsprechenden Zahlen 2400 und 2386. Hieraus folgt:  $\Phi^{(b)} < N \varepsilon_b \log_b N$ .

12. Der Übergang von 2 auf 3 und mehr Bahnen bringt für den Mischprozess keine Verbesserung in bezug auf die Zahl der Vergleichsoperationen. Diese Zahl ist jedoch nur *ein* Bestimmungsgrund für die Güte eines Sortierverfahrens. Neben diesem gibt es noch andere, die für die Dauer dieser Prozesse von Bedeutung sind, z. B. die *Zahl der Transfers*. Diese Zahl ist *proportional*  $\log_b N$ , wobei der Logarithmus zur Basis  $b =$  Zahl der Bahnen zu nehmen ist. Daraus ergibt sich eine wesentliche Reduktion mit steigendem  $b$ ; so ermässigt sich bei Übergang von 2 auf 3 Bahnen die Zahl der Transfers auf fast die Hälfte, was natürlich von praktischer Bedeutung ist. Wenn auch die Zahl der Vergleichsoperationen nicht abnimmt und das Programm länger und komplizierter wird, so wird man schon aus diesem Grunde das Mischen und Sortieren auf mehr als 2 Bahnen *bei grossen Datenmengen*, die externe Speicher benötigen, vorziehen, sofern es die technische Ausrüstung der elektronischen Anlage zulässt.

### Literaturverzeichnis

- [1] C. Boehm: Sortierprozesse auf elektronischen Rechenanlagen; in: Blätter der DGVM, Band V, Heft 2, S. 155–207, April 1961.
- [2] H. Nagler: An Estimation of the Relative Efficiency of two Internal Sorting Methods; in: Communications of the ACM, Vol. 3, p. 618–620, Nov. 1960.
- [3] H. H. Goldstine–J. v. Neumann: Planning and Coding of Problems for an Electronic Computing Instrument, Part II, Vol. II, p. 49 ff., Princeton, N. J., 1948.
- [4] C. Boehm: Die Bereinigung eines Versicherungsbestandes als Sonderfall des Mischprozesses; in: Blätter der DGVM, Band III, Heft 3, Oktober 1957, S. 323 ff.

Weitere Literaturangaben in [1].

## Résumé

L'auteur détermine l'espérance mathématique du nombre des opérations comparatives dans le cas d'un *interclassement sur 3 chemins* de  $N = 3n$  unités de triage, à l'aide de l'expression :

$$\varphi^{(3)} = \frac{2N^2}{N+3} \frac{5N+9}{6N+9}$$

Il tire de cette formule comme espérance mathématique du nombre des opérations comparatives pour le *trriage d'un interclassement itérative sur 3 chemins* (straight merging sort on 3 ways) lorsque  $N = 3^k$  ( $k \geq 5$ )

$$\Phi^{(3)} = 2N(3,475873 + (\log_3 N - 5)R)$$

$$0,83 < R < 0,8\bar{3}.$$

On obtient ainsi un nombre de comparaisons par unité plus *grand* qu'avec un procédé de triage sur 2 chemins.

Il est surprenant de constater que l'interclassement sur 2 chemins de  $n$  et  $n$  unités tout d'abord et de  $2n$  et  $n$  unités ensuite conduit à la *même* espérance mathématique que l'interclassement sur 3 chemins. Ce fait permet de fixer, pour le nombre des opérations comparatives en cas de interclassement sur plus de 3 chemins, des limites supérieures utilisables.

## Summary

The author computes the mean of the number of comparative operations by *merging on 3 ways* with  $N = 3n$  sorting units and finds

$$\varphi^{(3)} = \frac{2N^2}{N+3} \frac{5N+9}{6N+9}$$

From this formula he derives the mean of the number of comparative operations for the *straight merging sort on 3 ways*,  $N = 3^k$  ( $k \geq 5$ )

$$\Phi^{(3)} = 2N(3,475873 + (\log_3 N - 5)R)$$

$$0,83 < R < 0,8\bar{3}.$$

This leads to a *greater* number of comparisons per unit than by the sorting procedure on 2 ways.

It is surprising that the merging on 2 ways of firstly  $n$  and  $n$  units and then of  $2n$  and  $n$  units leads to the *same* expectation as the merging on 3 ways. This renders possible the definition of useful upper limits for the number of the comparative operations by merging on more than 3 ways.

### Riassunto

L'autore determina la speranza matematica del numero delle operazioni comparative nel *mescolare su 3 traettorie*  $N = 3n$  unità di assortimento per mezzo dell'espressione:

$$\varphi^{(3)} = \frac{2N^2}{N+3} \frac{5N+9}{6N+9}$$

e ricava da questa formula come speranza matematica del numero delle operazioni comparative per l'*assortimento dal mescolare continuo su 3 traettorie* (straight merging sort on 3 ways) quando  $N = 3^k$  ( $k \geq 5$ )

$$\Phi^{(3)} = 2N(3,475873 + (\log_3 N - 5)R)$$
$$0,83 < R < 0,8\bar{3}.$$

Ne risulta così un numero *maggiore* di comparazioni per unità che nel processo di assortimento su 2 traettorie.

È sorprendente constatare che il mescolare su 2 traettorie di  $n$  e  $n$  unità in primo luogo e in seguito di  $2n$  e  $n$  unità conduce alla *stessa* speranza matematica che il mescolare su 3 traettorie. Ciò permette di fissare limiti superiori utilizzabili per il numero di operazioni comparative nel mescolare su più di 3 traettorie.

