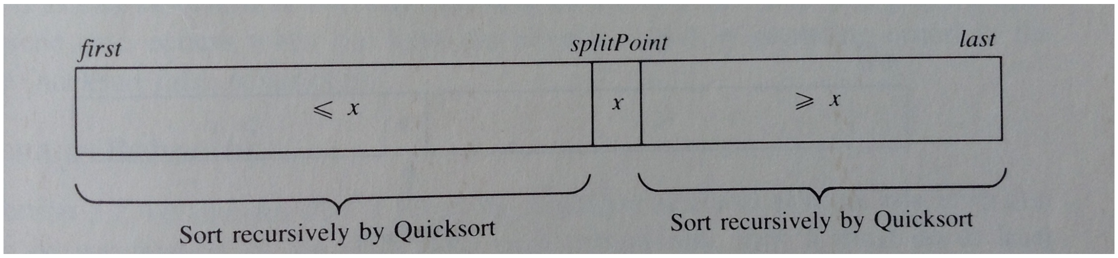
Solving a recurrence relation for the average – case running time A[n] for QuickSort :



$$A[n\_] := n - 1 + \frac{1}{n} \sum_{i=0}^{n-1} \left(A[i] + A[n - 1 - i]\right);$$

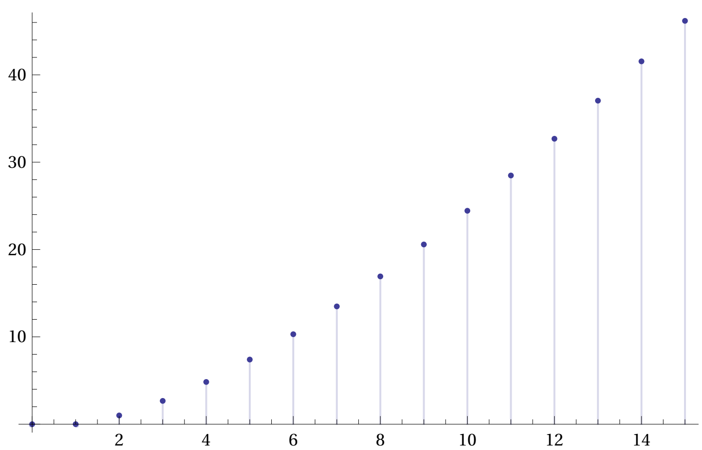
A[0] := 0;

A[1] := 0;

Mathematica cannot directly solve it but it can graph it.

(* The graphing is very slow because Mathematica computes values A[n] much more wastefully then Fibonacci numbers from the recurrence relation F[n]=F[n-1]+F[n-2]. *)

DiscretePlot[A[n], {n, 0, 15}, PlotTheme → "Classic"]



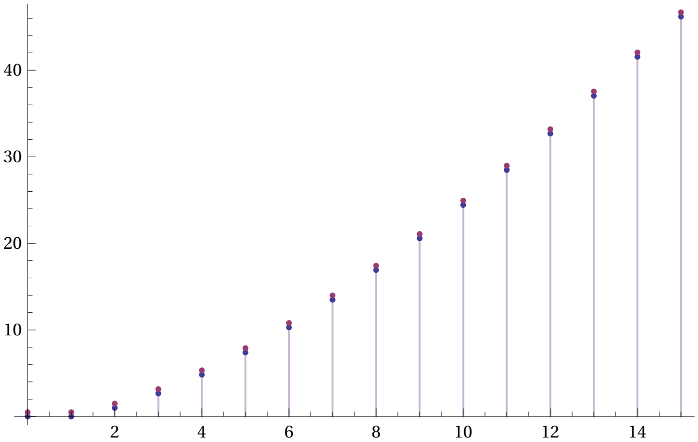
It may be verified by direct calculation that the function

$$A[n] = 2\,(n + 1) \left(\sum_{i=1}^{n} \frac{1}{i}\right) - 4\,n$$

satisfies the above recurrence relation for every $n \geq 1$.

The graphs of both solutions coincide (the second was plotted with extra + .5 added to it so that one can see two graphs) :

```
DiscretePlot[{A[n], 2 (n+1) (Sum[1/i, {i, 1, n}]) - 4 n + .5}, {n, 0, 15}, PlotTheme → "Classic"]
```



In order to prove that

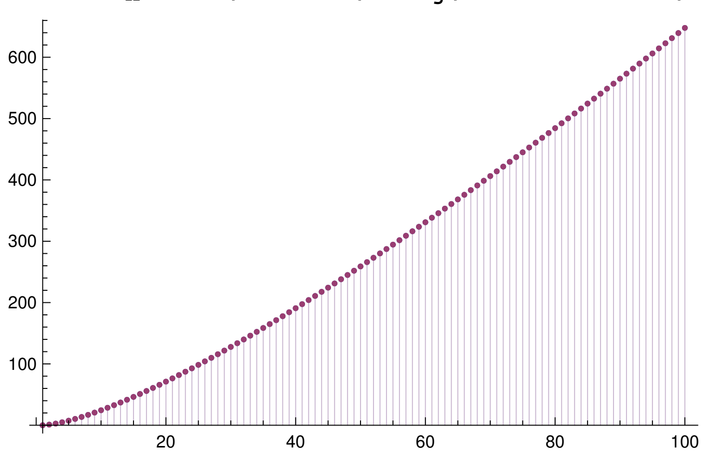
$$A[n] = 2\,(n+1)\left(\sum_{j=1}^{n}\frac{1}{j}\right) - 4\,n$$

is a solution of the above recurrence relation,
it suffices to show that for every n ≥ 1 :

$$2\,(n+1)\left(\sum_{j=1}^{n}\frac{1}{j}\right) - 4\,n ==$$
$$n - 1 + \frac{1}{n}\sum_{i=0}^{n-1}\left(2\,(i+1)\left(\sum_{j=1}^{i}\frac{1}{j}\right) - 4\,i + 2\,(n-i)\left(\sum_{j=1}^{n-1-i}\frac{1}{j}\right) - 4\,(n-1-i)\right)$$

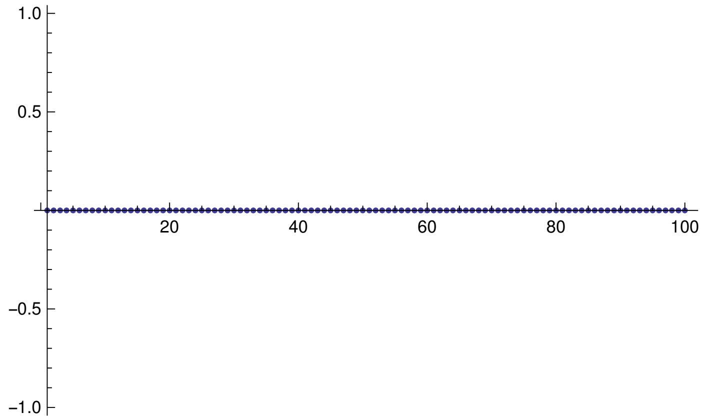
Here is an experimental verification by graphing

```
DiscretePlot[{ 2 (n+1) (Sum[1/j, {j, 1, n}]) - 4 n,
  n - 1 + 1/n Sum[(2 (i+1) (Sum[1/j, {j, 1, i}]) - 4 i + 2 (n-i) (Sum[1/j, {j, 1, n-1-i}]) - 4 (n-1-i)), {i, 0, n-1}]}, {n, 1, 100}]
```



The difference between the sides of the equation is 0

DiscretePlot$\left[\left\{2\,(n+1)\left(\sum_{j=1}^{n}\frac{1}{j}\right)-4\,n-\left(n-1+\frac{1}{n}\sum_{i=0}^{n-1}\left(2\,(i+1)\left(\sum_{j=1}^{i}\frac{1}{j}\right)-4\,i+2\,(n-i)\left(\sum_{j=1}^{n-1-i}\frac{1}{j}\right)-4\,(n-1-i)\right)\right)\right\},\ \{n,\ 1,\ 100\}\right]$



Here is an algebraic proof

In order to simplify calculations, we will eliminalte long summation $\sum_{i=0}^{n-1}$ from the recurrence relation. This can be obtained by calculating $n\,A[n]-(n-1)\,A[n-1]$.

This way the following equality may be obtained from the recurrence ralation for A[n] - see textbook for a straightforward derivation.

For every $n \geq 1$,

$$A[n]=(n+1)\,\frac{A[n-1]}{n}+\frac{2\,(n-1)}{n}$$

It suffices to show that our solution given by

$$A\,[n]=2\,(n+1)\left(\sum_{i=1}^{n}\frac{1}{i}\right)-4\,n$$

and, therefore,

$$A[n-1]=2\,n\left(\sum_{i=1}^{n-1}\frac{1}{i}\right)-4\,(n-1)$$

satisfies the equation.

Simplification of the right - hand side of the equation yields :

Expand$\left[(n+1)\,\frac{1}{n}\left(2\,n\left(\sum_{i=1}^{n-1}\frac{1}{i}\right)-4\,(n-1)\right)+\frac{2\,(n-1)}{n}\right]$

$$2+\frac{2}{n}-4\,n+2\,\text{HarmonicNumber}[-1+n]+2\,n\,\text{HarmonicNumber}[-1+n]$$

(* Note. HarmonicNumber$[k]=\sum_{i=1}^{k}\frac{1}{i}$ *)

$$\frac{2\,n}{n}+\frac{2}{n}-4\,n+2\left(\sum_{i=1}^{n-1}\frac{1}{i}\right)+2\,n\left(\sum_{i=1}^{n-1}\frac{1}{i}\right)$$

$$-4\,n+2\left(\sum_{i=1}^{n}\frac{1}{i}\right)+2\,n\left(\sum_{i=1}^{n}\frac{1}{i}\right)$$

The left - hand side of the equation is

$\text{Expand}\left[2\,(n+1)\left(\sum_{i=1}^{n}\frac{1}{i}\right)-4\,n\right]$

$-4\,n+2\,\text{HarmonicNumber}[n]+2\,n\,\text{HarmonicNumber}[n]$

$-4\,n+2\left(\sum_{i=1}^{n}\frac{1}{i}\right)+2\,n\left(\sum_{i=1}^{n}\frac{1}{i}\right)$

So, lhs = rhs.

Bingo !

End of algebraic solution

Substituting the Euler ' s approximation of $\sum_{i=1}^{n}\frac{1}{i}$ in the above formula (see file Summationa.nb) we obtain :

$$A[n] \approx 2\,(n+1)\left(\text{Log}[n]+.577+\frac{1}{2\,n}\right)-4\,n$$

or

$\text{Simplify}\left[2\,(n+1)\left(\text{Log}[n]+.577+\frac{1}{2\,n}\right)-4\,n\right]$

$2.154+\frac{1}{n}-2.846\,n+2\,(1+n)\,\text{Log}[n]$

Since Log[n] = Log[2] × Log2[n], we conclude :

$$A[n] \approx 2.154` + \frac{1}{n} - 2.846`\,n + 2\,\text{Log}[2]\,(1+n)\,\text{Log2}[n]$$

and by

N[2 Log[2]]

1.3862943611198906`

$$A[n] \approx 2.154+\frac{1}{n}-2.846\,n+1.386\,(1+n)\,\text{Log2}[n]$$

or

$$A[n] \approx 1.386\,(1+n)\,\text{Log2}[n]-2.846\,n+2.154+\frac{1}{n}$$

The difference between the left - hand side and the right - hand side of the above approximation converges to 0 faster than $\frac{1}{n}$.

Using closer approximation Log[n] + .577 + $\frac{1}{2\,n}$ - $\frac{1}{12\,n^2}$ of Euler's formula one can get

A[n] ≈

$$1.3862943611198906`\,(n+1)\,\text{Log2}[n] - 2.8455686701969345`\,n + 2.1544313298030655` + \frac{1}{n}\,0.8333333333333334` - (0.16666666666666666`/n^2)$$

the difference between which and the exact A[n] converges to 0 faster than $\frac{1}{n^2}$.

The part highlighted yellow and its derivation is mandatory for all students; the rest of the aproximation and its derivation is optional - they are shown here for illustration purposes only.

(The textbook approximation was

A[n] ≈ 1.386 n Log2[n] - 2.846 n

- good enough if you don't want to be too accurate, although

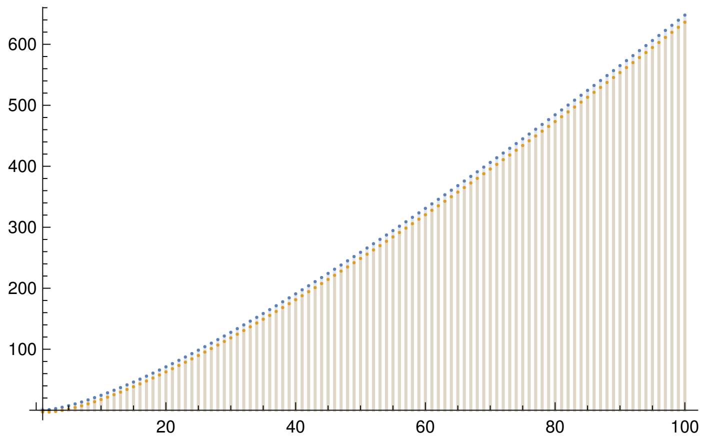
A[n] ≈ 1.386 (n + 1) Log2[n] - 2.846 n

is somewhat better - see the 2 plots below.)

Textbook approximation vs the exact value :

In[1]:= DiscretePlot$\left[\left\{2\ (n+1)\left(\sum_{j=1}^{n}\frac{1}{j}\right)-4\ n,\ 1.386\ n\ \mathrm{Log2}[n]-2.846\ n\right\},\ \{n,\ 1,\ 100\}\right]$

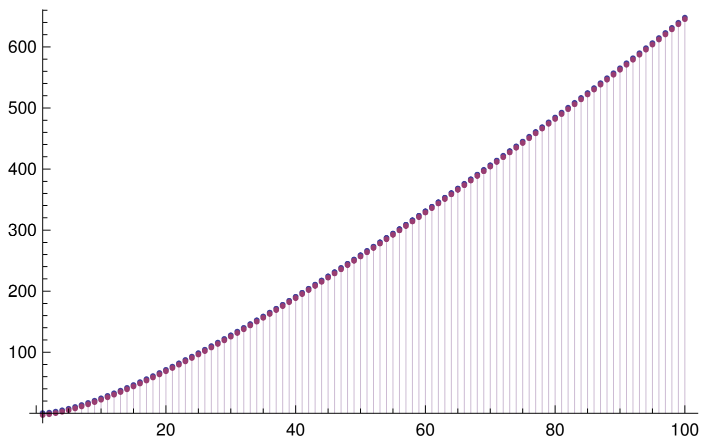
Out[1]=



The difference between the exact value and textbook approximation diverges to ∞ with n.

Improved approximation vs the exact value :

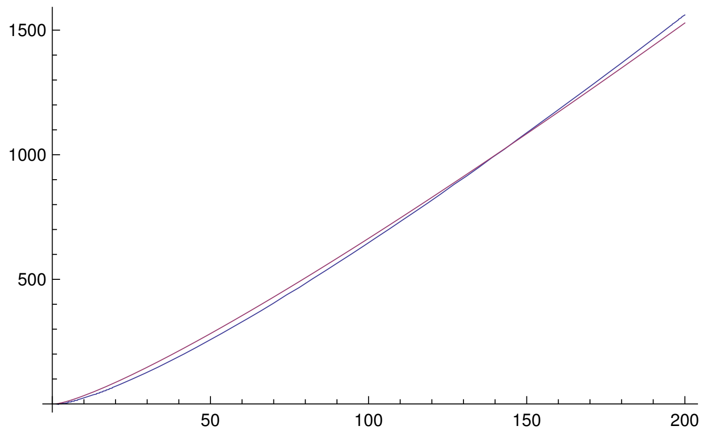
DiscretePlot$\left[\left\{2\ (n+1)\left(\sum_{j=1}^{n}\frac{1}{j}\right)-4\ n,\ 1.386\ (n+1)\ \mathrm{Log2}[n]-2.846\ n\right\},\ \{n,\ 1,\ 100\}\right]$



$2\ (n+1)\left(\sum_{j=1}^{n}\frac{1}{j}\right)-4\ n < n \lg n$ for $n \le 141$ and

$2\ (n+1)\left(\sum_{j=1}^{n}\frac{1}{j}\right)-4\ n > n \lg n$ for $n \ge 142$

```
Plot[{2 (n + 1) (∑_{j=1}^{n} 1/j) - 4 n, n Log2[n] }, {n, 1, 200}]
```




```
Table[{N[2 (n + 1) (∑_{j=1}^{n} 1/j) - 4 n], N[n Log2[n]] }, {n, 141, 142}]
{{1006.38, 1006.68}, {1015.46, 1015.26}}
```

The difference between the exact value and improved approximation converges to about 2.154 as n → ∞.

Check of the more precise approximation :

```
Simplify[2 (n + 1) (Log[n] + .577 + 1/(2 n) - 1/(12 n^2)) - 4 n]
```

$$2.154 - \frac{0.166667}{n^2} + \frac{0.833333}{n} - 2.846\, n + 2\,(1 + n)\,\mathrm{Log}[n]$$

```
DiscretePlot[{2 (n + 1) (∑_{j=1}^{n} 1/j) - 4 n, 2.154` - 0.16666666666666666` / n^2 +
   0.8333333333333334` / n - 2.846` n + 2 (1 + n) Log[n]}, {n, 1, 100}]
```