Asymptotics of linear divide-and-conquer recurrences

Philippe Dumas
Résumé par Bruno Salvy

Séminaire de Combinatoire Philippe Flajolet
Institut Henri Poincaré, Séance du 6 juin 2013

Abstract

Asymptotics of divide-and-conquer recurrences is usually dealt either with elementary inequalities or with sophisticated methods coming from analytic number theory. Philippe Dumas proposes a new approach based on linear algebra. The example of the complexity of Karatsuba's algorithm is used as a guide in this summary.

The complexity analysis of divide-and-conquer algorithms gives rise to recurrences that relate the cost at size n to the cost at fractions of n. For instance, the complexity of Karatsuba's multiplication algorithm for polynomials of degree n is governed by

$$c(n) = n + 3c\left(\lceil n/2 \rceil\right). \tag{1}$$

The first values taken by this sequence with $c(1) = 1$ are displayed in Figure 1.

The linear term n is of course dependent on the complexity model. However, the analysis is quite robust and any function growing linearly would lend itself to this analysis, leading to similar results with minor technical adjustments.

Notation In this presentation, general consideration are interlaced with this particular example. In order to help distinguish between the general and the specific, we use blue characters to display the particulars of the example.

1 Divide-and-conquer Recurrences

A more complicated example is provided by the analysis of a recent "dichopile" algorithm due to J. Oudinet. This algorithm performs random generation from a regular language with uniform

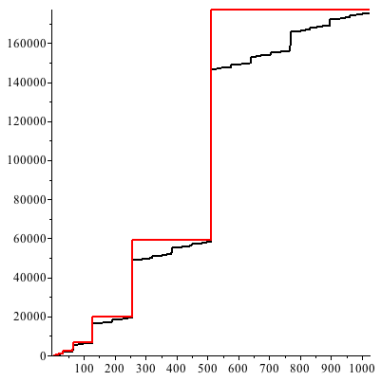

Figure 1: The sequence $c(n)$ (black) and its upper bound from Section 3 (red)

distribution over its words of length n. It is designed to avoid using too much memory by storing intermediate values for sizes $n/2, 3n/4, 7n/8, \ldots$. Its complexity for length n (in terms of the number of arithmetic operations) obeys the recurrence

$$f(n) = n + f(\lfloor n/2 \rfloor - 1) + g(\lceil n/2 \rceil), \quad \text{where} \quad g(n) = f(\lfloor n/2 \rfloor - 1) + g(\lceil n/2 \rceil), \quad f(1) = 1, g(1) = 0.$$

The classical Rudin-Shapiro sequence also satisfies a divide-and-conquer recurrence: it is the sequence of coefficients of the polynomial of degree n with coefficients in $\{-1, 1\}$ whose maximum modulus over $|z| = 1$ is mimimal (normalized with $P(0) = 1$). It turns out that whatever the degree, this sequence satisfies

$$a_{2n} = a_n, \qquad a_{2n+1} = (-1)^n a_n.$$

2 2-rational sequences

The Rudin-Shapiro sequence is a *2-rational* sequence, which means that the vector space spanned by the sequence (a_n) and the iterates of the operators $S_0 : (u_n) \mapsto (u_{2n})$ and $S_1 : (u_n) \mapsto (u_{2n+1})$ applied to it is finite dimensional. Indeed, the identities

$$S_0(a_n) = (a_n), \quad S_0(S_1(a_n)) = (a_{4n+1}) = (a_n) \quad \text{and} \quad S_1(S_1(a_n)) = (a_{4n+3}) = (-a_{2n+1})$$

show that all these sequences are generated by (a_n) and $S_1(a_n)$. Similarly, the sequence $(c(n))$ from the analysis of Karatsuba's algorithm is 2-rational, with a space of dimension 4 generated by $(1), (n), (c(n)), (c(n+1))$:

$$\alpha + \beta n + \gamma c(n) + \delta c(n+1) \overset{S_0}{\mapsto} \alpha + \beta(2n) + \gamma(2n + 3c(n)) + \delta(2n + 1 + 3c(n+1)), \tag{2}$$

$$\alpha + \beta n + (\gamma + \delta)c(n) + \delta\,(c(n+1) - c(n)) \overset{S_1}{\mapsto} \alpha + \beta(2n+1) + (\gamma + \delta)(2n + 1 + 3c(n+1)) \tag{3}$$
$$+\delta(\underbrace{(2(n+1) + 3c(n+1)) - (2n + 1 + 3c(n+1))}_{1}).$$

A similar but longer computation shows that the sequence $f(n)$ in the cost of the dichopile algorithm is also 2-rational (the dimension is 7).

A 2-rational sequence $(c(n))$ can be given by a *linear representation*, i.e., matrices A_0 and A_1 of the operators S_0 and S_1, the vector L of initial values of the elements of the corresponding basis at $n = 1$, and the vector C giving the coordinates of $(c(n))$. In Karatsuba's example, the translation of Equations (2-3) yields

$$L = \begin{pmatrix} 1 & 1 & 1 & 5 \end{pmatrix}, \quad A_0 = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 2 & 2 & 2 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

From this representation, the value of the sequence at n, written in base 2 as $n = (1 b_k \ldots b_1 b_0)_2$, is given by the product

$$c(n) = L A_{b_k} \cdots A_{b_1} A_{b_0} C.$$

This shows how 2-rational sequences generalize rational sequences (like Fibonacci's sequence). In both cases, rational series from the theory of formal languages underly the structure and the nth element can be computed by $O(\log n)$ matrix products.

Bounding the norms shows that such a sequence grows at most like $O(M^{\log_2 n})$ for some M that can be taken as $\max(\|A_0\|, \|A_1\|)$. This also shows that not all solutions of divide-and-conquer recurrences are 2-rational. For instance, partitions of an integer n into powers of 2 are obtained either by adding 1 to a similar partition of $n-1$ or, if n is even, by multiplying by 2 all parts of a partition of $n/2$. Thus, their number satisfies a system that looks similar to the previous ones: $b_{2n+1} = b_{2n} + 1$, $b_{2n} = b_n + b_{2n-1}$. However, Mahler showed [9] that b_{2n} behaves asymptotically like $\exp(\ln^2 n/2)$, which proves that it cannot be 2-rational.

3 Elementary inequalities

In a large number of cases, simple bounds can be obtained rather easily. For instance, the sequence $c(n)$ from Karatsuba's complexity equation (1) is increasing (by induction) and therefore it is sufficient to consider it at powers of 2 and bound the value at any n by the value at $2^{\lceil \log_2 n \rceil} < 2n$. If $n = 2^k$ and $c(n) =: d(k)$, the equation becomes

$$d(k) = 2^k + 3d(k-1) = 2^k + 3 \cdot 2^{k-1} + d(k-2) = \cdots = 3^k(1 + \frac{2}{3} + \frac{4}{9} + \dots) \leq 3^{k+1} = 3 \cdot n^{\log_2 3},$$

which gives the correct order of growth for the sequence. (See Figure 1).

This technique is often convenient in rough complexity analyses. Its first version is due to Bentley, Haken and Saxe in 1980 [2]. Modern versions appear under the name "Master Theorem" in Cormen *et alii*'s *Introduction to Algorithms* [3]. More elaborate variants have been developed. An easy-to-use and relatively general one has been proposed by Yap recently [10].

4 Perron's formula

Bounding only at powers of 2 misses the fine behavior of the algorithm (see Figure 1). Another route uses the Dirichlet series

$$C(s) = \sum_{k \geq 1} \frac{c(k+2) - c(k+1)}{k^s}$$

and Perron's formula (presented for instance in Apostol's book [1, Th. 11.18])

$$c(n) = c(2) + \lim_{T \to \infty} \frac{1}{2\pi i} \int_{\sigma - iT}^{\sigma + iT} C(s) \frac{(n - 3/2)^s}{s} \, ds.$$

In the case of Karatsuba's sequence, the recurrence equation (1) translates into

$$\begin{aligned} C(s) &= \zeta(s) + 3 \cdot 2^{-s} C(s) + 12 + 3 \sum_{k \geq 1} (c(k+2) - c(k+1)) \left(\frac{1}{(2k+1)^s} - \frac{1}{(2k)^s} \right) \\ &=: \zeta(s) + 3 \cdot 2^{-s} C(s) + G(s), \end{aligned}$$

where ζ is Riemann ζ-function and G has abscissa of convergence smaller than that of C. This equation rewrites

$$C(s) = \frac{\zeta(s) + G(s)}{1 - 3 \cdot 2^{-s}}.$$

In general, the next stage of the analysis consists in shifting the vertical line of integration past the right-most poles of $C(s)$, that are here vertically aligned at $(\log 3 + 2k\pi i)/\log 2$, $k \in \mathbb{Z}$, and picking up residues there. Serious analytic precautions have to be taken in order to ensure

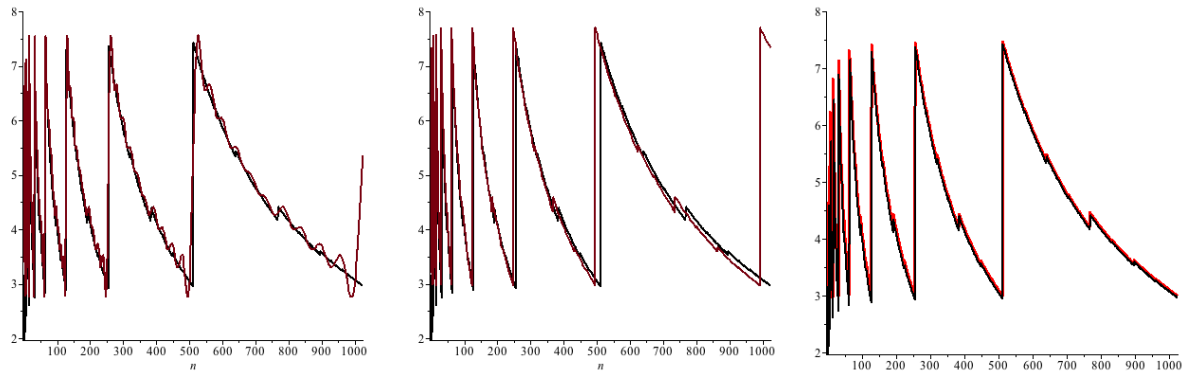

Figure 2: Sequence $c(n)/n^{\log_2 3}$ (black) vs asymptotic behavior (red). Left: Perron's formula with the sum of the first 10 terms of the Fourier series. Middle: Perron's formula with the sum of the first 40 terms of Eq. (4). Right: the linear algebra asymptotics of Eq. (7).

convergence (or even mere meaningfulness) of the result. In our example, we thus obtain the following first terms of a Fourier series:

$$c(n)n^{-\log_2 3} \approx 4.856 + .261\cos(2\pi\log_2(n)) + 1.308\sin(2\pi\log_2(n)) \\ + 0.055\cos(4\pi\log_2(n)) + .712\sin(4\pi\log_2(n)) + \cdots .$$

Each coefficient requires summing a rather large number of the series G and a Gibbs phenomenon can be observed (see Figure 2).

Another way of obtaining this periodic function is to expand the numerator $\zeta(s)+G(s)$ as a Dirichlet series and integrate term by term, using the fact [8] that for $x > 1$ such that $\log_2 x \notin \mathbb{Z}$ and for $c > \log_2 3$,

$$\frac{1}{2\pi i}\lim_{T\to\infty}\int_{c-iT}^{c+iT}\frac{x^s}{1-3\cdot 2^{-s}}\frac{ds}{s} = \frac{3^{\lfloor\log_2 x\rfloor+1}-1}{2}.$$

This gives an asymptotic expansion of the form

$$n^{\log_2 3}\Psi(\log_2 n) + O(n^{1+\epsilon}),$$

where Ψ is a periodic function with period 1, given by an expansion that is very easy to compute:

$$\Psi(x) = \frac{39}{2}\cdot 3^{\lfloor x\rfloor - x} - 57\cdot 3^{\lfloor x-\ln_2 2\rfloor - x} + 60\cdot 3^{\lfloor x-\ln_2 3\rfloor - x} - 3\cdot 3^{\lfloor x-\ln_2 4\rfloor - x} + 6\cdot 3^{\lfloor x-\ln_2 5\rfloor - x} + \cdots . \quad (4)$$

A relatively friendly theorem based on the use of Perron's formula has been given recently by Drmota and Szpankowski [5].

5 Linear Algebra Approach

Recently, Philippe Dumas has shown that in the case of 2-rational sequences, the oscillatory part of the asymptotic behavior turns out to be accessible through linear algebra considerations [7].

The starting point is to consider the sum of all the values of the sequence up to n:

$$s_n = LC + LA_0C + LA_1C + LA_0A_0C + LA_0A_1C + LA_1A_0C + \cdots + LA_{b_k}\cdots A_{b_0}C, \quad (5)$$

where again n is given by its binary expansion $(1b_k\cdots b_0)_2$. There is no loss of generality in considering a sum, since the difference of a 2-rational sequence is 2-rational as well and its sum

is the original sequence. In our running example, it is not difficult to check that $c(n+1)-c(n)$ is defined by

$$L = \begin{pmatrix} 1 & 4 \end{pmatrix}, \quad A_0 = \begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The sum (5) decomposes as a part where all binary words of length $0, 1, \ldots, k$ over $\{A_0, A_1\}$ are present, plus a part whose length depends on n, where the oscillatory behavior takes place. This is captured by introducing the matrix function

$$S_k(x) := \sum_{\substack{(0.w)_2 \le x \\ w \in \{0,1\}^k}} A_w,$$

so that the sum of interest becomes

$$s_n = L(I_m + Q + \cdots + Q^{k-1})C + LS_k\left(\frac{n-2^k}{2^k}\right)C, \quad \text{with } k = \lfloor \log_2 n \rfloor,$$

where $Q = A_0 + A_1$ and I_m is the identity matrix (its dimension m is that of the representation). The asymptotic behavior of this sum is governed by the largest eigenvalue of Q and the *joint spectral radius* of A_0 and A_1, which is, by definition,

$$\rho_\star = \lim_{k\to\infty} \max_{|w|=k} \{\|A_{w_1} \cdots A_{w_k}\|^{1/k}\}.$$

When the largest eigenvalue of Q has modulus not smaller than $\rho_\star$, it dictates the asymptotic behavior of s_n. In our example, they both turn out to be equal to 3, which slightly complicates the analysis.

Recurrence Equations First, distinguishing the binary words of length $k+1$ according to their first letter leads to a recurrence equation:

$$S_{k+1}(x) = \begin{cases} A_0 \cdot S_k(2x), & \text{if } x < 1/2, \\ A_0 \cdot Q^k + A_1 \cdot S_k(2x-1), & \text{otherwise.} \end{cases}$$

Eventually, we are interested in the product $S_k(x) \cdot C$, so that it is useful to consider the decomposition of C along the eigenspaces of Q. In particular, subspaces corresponding to eigenvalues of modulus smaller than $\rho^\star$ contribute to error terms in the expansion.

In our example, there are two eigenvectors V and W, and all the relevant information is given by

$$\begin{cases} QV = 3V, \\ QW = 2W, \end{cases} \quad \begin{cases} A_0V = 3V - 3W, \\ A_0W = W, \end{cases} \quad \begin{cases} A_1V = 3W, \\ A_1W = W, \end{cases} \quad C = V - 2W.$$

An easy induction shows that the coordinate of $S_k(x)C$ on V is exactly 3^k. It is then natural to define a function $\phi_k(x)$ by

$$S_k(x)C = 3^k V + 3^k \phi_k(x) W.$$

The recurrence over $S_k(x)$ then translates into

$$\phi_{k+1}(x) = \begin{cases} -1 + \frac{1}{3}\phi_k(2x), & \text{if } x < 1/2, \\ -\left(\frac{2}{3}\right)^{k+1} + \frac{1}{3}\phi_k(2x-1), & \text{otherwise.} \end{cases}$$

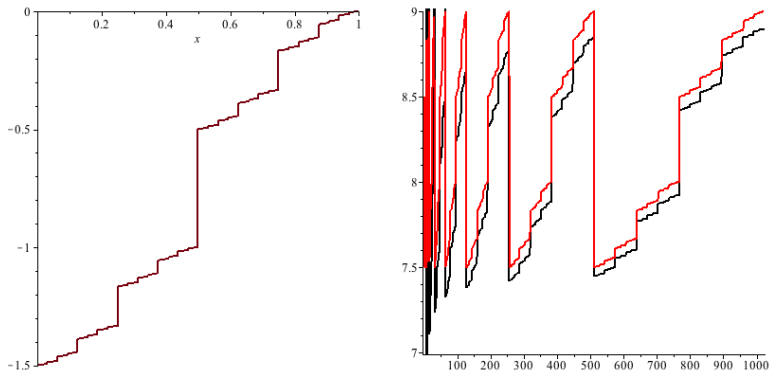

Figure 3: Left: the solution of the dilation equation (6). Right: the sequence $c(n+1)/3^{\lfloor \log_2 n \rfloor}$ (black) vs its asymptotic behavior by Eq. (7) (red).

Dilation Equation Letting k tend to infinity, the equation above becomes a simple functional equation

$$\phi(x) = \begin{cases} -1 + \frac{1}{3}\phi(2x), & \text{if } x < 1/2, \\ \frac{1}{3}\phi(2x-1), & \text{otherwise.} \end{cases} \tag{6}$$

Setting for simplicity $\phi(x) = -\frac{3}{2} + \frac{3}{2}\psi(x)$, this turns into the simpler looking:

$$\psi(x) = \frac{1}{3}\psi(2x) + \frac{1}{3}\psi(2x-1), \qquad \text{with } \psi(x) = 1 \text{ for } x > 1 \text{ and } \psi(x) = 0 \text{ for } x < 0.$$

This is a fixed point equation in the space of functions from $\mathbb{R}^+$ to $[0,1]$, space which is complete for the supremum norm. Thus it has a solution, which is unique. It can be computed by the "cascade algorithm" [4, §6.5] that consists simply in recursing sufficiently many times. The resulting (discontinuous) function ϕ is displayed in Figure 3. It turns out to be closely related to Cantor's "devil staircase".

Convergence Letting ϕ denote the solution of the dilation equation (6), we then define the deviation $e_k(x)$ of ϕ_k from ϕ by $\phi_k(x) = \phi(x) + e_k(x)$. This function satisfies a recurrence equation deduced from the previous ones.

In our example, we get

$$e_{k+1}(x) = \begin{cases} \frac{1}{3}e_k(2x), & \text{if } x < 1/2, \\ -\left(\frac{2}{3}\right)^{k+1} + \frac{1}{3}e_k(2x-1), & \text{otherwise.} \end{cases}$$

For $k = 0$ we have $S_0(x)C = C = V - 2W$ and since $\|\phi\|_\infty \le 3/2$ we deduce $\|e_0\|_\infty \le 7/2$. From there, an easy induction using the recurrence shows that $\|e_k\|_\infty \le \frac{7}{2}\left(\frac{2}{3}\right)^k \to 0$.

Asymptotic Expansion Multiplying by L on the left then yields a very explicit asymptotic expansion:

$$\begin{aligned} c(n+1) - c(1) = s_n &= L(\mathrm{Id} + Q + \cdots + Q^{k-1})C + LS_k\left(\frac{n-2^k}{2^k}\right)C, \quad \text{with } k = \lfloor \log_2 n \rfloor \\ &= L\left(\frac{3^k-1}{2}V + O(2^k)W\right) + L(3^kV + 3^k\phi(2^{\{\log_2 n\}}-1)W + O(2^k)W), \\ &= 3^k\left(9 + \phi(2^{\{\log_2 n\}}-1)\right) + O(2^k), \\ c(n+1) &= 3^{\lfloor \log_2 n \rfloor}\left(9 + \phi(2^{\{\log_2 n\}}-1)\right) + O(n). \end{aligned} \tag{7}$$

The match with $c(n)/n^{\log_2 3}$ is better than with the previous method (see Figure 2) and the normalization by $3^{\lfloor \log_2 n \rfloor}$ reveals more clearly the self-similarity of the curve (Figure 3).

6 A general theorem

The example of Karatsuba's algorithm is very special because the spectral radius of the matrix Q is equal to the joint spectral radius of A_0 and A_1. In general, the situation becomes a bit simpler and a complete characterisation of the asymptotic behaviour can be given.

Theorem 1 (Dumas 2013)**.** *If (s_n) is a sequence whose backward differences is defined by a linear representation L, A_0, A_1, C, then it behaves asymptotically like*

$$s_N = \sum_{\rho,\theta,m} n^{\log_2 \rho}\binom{\log_2 n}{m} e^{i\theta \log_2 n}\Phi_{\rho,\theta,m}(\log_2 n) + O(n^{\log_2 r}),$$

where the sum is over eigenvalues ρ of Q such that $|\rho| > r > \rho_\star$, m is a nonnegative integer bounded by the multiplicity of the eigenvalue and Φ is a periodic function with period 1.

7 Extensions

By focusing on one example in this summary, we have left out several other issues of the general case. We refer to the original articles [6, 7] for more information, including the case of B-rational sequences with $B \neq 2$ and a full treatment of the dichopile example.

References

[1] Apostol, T. M. *Introduction to Analytic Number Theory.* Springer-Verlag, 1976.

[2] Bentley, J. L., Haken, D., and Saxe, J. B. A general method for solving divide-and-conquer recurrences. *SIGACT News 12*, 3 (Sept. 1980), 36–44.

[3] Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C. *Introduction to algorithms*, third ed. MIT Press, Cambridge, MA, 2009.

[4] Daubechies, I. *Ten lectures on wavelets*, vol. 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[5] Drmota, M., and Szpankowski, W. A master theorem for discrete divide and conquer recurrences. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms* (Philadelphia, PA, 2011), SIAM, pp. 342–361.

[6] Dumas, P. Joint spectral radius, dilation equations, and asymptotic behavior of radix-rational sequences. *Linear Algebra Appl. 438*, 5 (2013), 2107–2126.

[7] Dumas, P. Rational series and asymptotic expansion for linear homogeneous divide-and-conquer recurrences. April 2013.

[8] Hardy, G. H., and Riesz, M. *The General Theory of Dirichlet's Series.* Cambridge Tracts in Mathematics. Cambridge University Press, 1915.

[9] Mahler, K. On a special functional equation. *Journal of the London Mathematical Society 15* (1940), 115–123.

[10] Yap, C. A real elementary approach to the master recurrence and generalizations. In *Theory and applications of models of computation*, vol. 6648 of *Lecture Notes in Comput. Sci.* Springer, Heidelberg, 2011, pp. 14–26.