The q-Catalan numbers: A saddle point approach

Guy Louchard* and Helmut Prodinger†

January 9, 2008

Abstract

Using the saddle point method, we obtain from the generating function of the q-Catalan numbers and Cauchy's integral formula asymptotic results in central and non-central regions.

Keywords: q-Catalan numbers, saddle point method.

1 Introduction

The q-Catalan numbers $C_n(q)$ are defined as

$$C_n(q) = \frac{1-q}{1-q^{n+1}} \frac{(q;q)_{2n}}{(q;q)_n (q;q)_{2n}},$$

with $(q;q)_n = (1-q)(1-q^2)\dots(1-q^n)$. Note that $C_n = C_n(1)$, a Catalan number, and the polynomial $F_n(q) = C_n(q)/C_n$ is the probability generating function of a distribution X_n that we call the *Catalan* distribution. Is was shown recently in [1] that this distribution is asymptotically normal.

In this paper, we want to sharpen and extend this result using the saddle point method. We follow our own paper [4] where we treated the random variable "number of inversions of a permutation" whose probability generating function is of a related type. We will get explicit error terms for the approximation to the normal distribution in the central region, and are able to improve the approximation by including as many terms as we want. We also investigate the non-central region near n. Of course, we must rely on a computer algebra system (Maple).

Since we attack $F_n(q)$ and $C_n(q)$ mostly with analytic methods, we find it more appropriate to replace the letter q by z; note that

$$C_n(z) = \prod_{i=1}^{n-1} \frac{1-z^{n+i+1}}{1-z^{i+1}}.$$

As mentioned in Chen et al. [1], it is easy to get the mean m and variance σ^2 of the random variable X_n characterized by

$$P_n(i) := \mathbb{P}(X_n = i) := \frac{[z^i]C_n(z)}{C_n},$$

namely

$$m = \frac{n(n-1)}{2},$$
$$\sigma^2 = \frac{n(n^2-1)}{6}.$$

*Université Libre de Bruxelles, Département d'Informatique, CP 212, Boulevard du Triomphe, B-1050 Bruxelles, Belgium, email: louchard@ulb.ac.be

†University of Stellenbosch, Mathematics Department, 7602 Stellenbosch, South Africa, email: hproding@sun.ac.za

(In the last section we sketch how these moments and higher ones can be computed quite easily.)

By Cauchy's theorem,

$$P_n(j) = \frac{1}{2\pi i}\int_{\mathcal{C}} F_n(z)\frac{dz}{z^{j+1}},$$

where $\mathcal{C}$ is, say, a circle around the origin passing (approximately) through the saddle point.

As general references for the application of the saddle point method in enumeration we cite [3, 5].

This method will provide local limit theorems, and we will get (potentially) as many lower order terms as we want, to make the approximation as accurate as we like.

The quality of the approximation for various choices of error terms will be displayed.

Our main results are the approximation (7) near the mean $n(n-1)/2$ and the approximation (16) near n:

Theorem 1.1 *Set $j = m + x\sigma$, with $m = n(n-1)/2$ and $\sigma = \sqrt{n(n^2-1)/6}$. Then, for fixed x, the following approximation holds:*

$$P_n(j) \sim e^{-x^2/2}\cdot \exp\left((-9/40 + 9/20x^2 - 3x^4/40)/n\right)\Big/(2\pi n^3/6)^{1/2}.$$

Theorem 1.2 *For fixed k, the following approximation holds:*

$$P_n(n-k) \sim 4^{-n}e^{\pi\sqrt{6n}/3}\frac{2^{-1/2}\pi^{3/2}}{12}\left(1 - \frac{2^{1/2}3^{1/2}(216+13\pi^2+24\pi^2 k)}{144\pi\sqrt{n}}\right).$$

Approximations near other values could also be obtained with our methods.

2 The Gaussian limit – approximation near the mean

We consider values $j = m + x\sigma$, in a neighbourhood of the mean m.

The Gaussian limit of X_n is obtained in [1]. Let us recover this result with the *saddle point method,* with an additional correction of order $1/n$. We know that

$$P_n(j) = \frac{1}{2\pi i}\int_{\Omega} e^{S(z)}dz$$

where

$$S(z) := \ln(F_n(z)) - (j+1)\ln z$$

Since $F_n(z)$ is a polynomial, the analyticity restriction can be ignored.

We split the exponent of the integrand $S = \ln(F_n(z)) - (j+1)\ln z$ as

$$\begin{aligned}
S &:= S_1 + S_2, && (1)\\
S_1 &:= \sum_{i=1}^{n-1}\left(\ln(1-z^{n+i+1}) - \ln(n+i+1) - \ln(1-z^{i+1}) + \ln(i+1)\right),\\
S_2 &:= -(j+1)\ln z.
\end{aligned}$$

Set

$$S^{(i)} := \frac{d^i S}{dz^i}.$$

To use the saddle point method, we must find the solution of

$$S^{(1)}(\tilde{z}) = 0. \qquad (2)$$

Set $\tilde{z} := z^* - \varepsilon$, where, here, $z^* = 1$. (This notation always means that z^* is the approximate saddle point and $\tilde{z}$ is the exact saddle point; they differ by a quantity that has to be computed to some degree of accuracy.) This leads, to first order, to

$$\Big(\frac{n^2}{2} - \frac{n}{2} - j - 1\Big) + \Big(-\frac{n^3}{6} + \frac{n^2}{2} - \frac{n}{3} - j - 1\Big)\varepsilon = 0. \tag{3}$$

Set $j = m + x\sigma$ in (3). This shows that, asymptotically, ε is given by a *Puiseux series* of powers of $n^{-1/2}$, starting with $-\sqrt{6}x/n^{3/2}$. To obtain the next terms, we compute the following terms in the expansion of (2), i.e., we first obtain

$$\Big(\frac{n^2}{2} - \frac{n}{2} - j - 1\Big) + \Big(-\frac{n^3}{6} + \frac{n^2}{2} - \frac{n}{3} - j - 1\Big)\varepsilon + \Big(-\frac{n^3}{3} + \frac{n^2}{2} - \frac{n}{4} - j - 1\Big)\varepsilon^2 = 0. \tag{4}$$

More generally, even powers ε^{2k} lead to a $\mathcal{O}(n^{2k+1})\cdot\varepsilon^{2k}$ term and odd powers ε^{2k+1} lead to a $\mathcal{O}(n^{2k+3})\cdot\varepsilon^{2k+1}$ term. Now we set $j = m + x\sigma$, expand into powers of $n^{-1/2}$ and equate each coefficient with 0. This leads successively to a full expansion of ε. Note that to obtain a given precision of ε, it is enough to compute a given finite number of terms in the generalization of (4). We obtain

$$\begin{aligned}\varepsilon = &-\sqrt{6}xn^{-3/2} - \frac{3\sqrt{6}}{10}n^{-5/2} + (-3x^2-6)n^{-3} \\ &+ x\sqrt{6}\Big(\frac{99}{700}x^5 - \frac{3}{10}x^3 - \frac{1}{2}x\Big)n^{-7/2} + \mathcal{O}(n^{-4}).\end{aligned}$$

We have, with $\tilde{z} := z^* - \varepsilon$,

$$P_n(j) = \frac{1}{2\pi i}\int_\Omega \exp\Big[S(\tilde{z}) + S^{(2)}(\tilde{z})(z-\tilde{z})^2/2! + \sum_{l=3}^\infty S^{(l)}(\tilde{z})(z-\tilde{z})^l/l!\Big]dz$$

(note carefully that the linear term vanishes). Set $z = \tilde{z} + i\tau$. This gives

$$P_n(j) = \frac{1}{2\pi}\exp[S(\tilde{z})]\int_{-\infty}^{\infty}\exp\Big[S^{(2)}(\tilde{z})(i\tau)^2/2! + \sum_{l=3}^\infty S^{(l)}(\tilde{z})(i\tau)^l/l!\Big]d\tau. \tag{5}$$

Let us first analyze $S(\tilde{z})$. We obtain

$$\begin{aligned}S_1(\tilde{z}) &= -\sqrt{6}x\sqrt{n}/2 + x^2/2 \\ &\quad + \sqrt{6}[3x^3/20 - x/2]/\sqrt{n} + [3 + 9x^4/40]/n + \mathcal{O}(n^{-3/2}), \\ S_2(\tilde{z}) &= -\sqrt{6}x\sqrt{n}/2 - x^2 \\ &\quad + \sqrt{6}[-3x^3/20 + x/2]/\sqrt{n} + [-3 - 3x^4/10]/n + \mathcal{O}(n^{-3/2}),\end{aligned}$$

and so

$$S(\tilde{z}) = -x^2/2 - \frac{3x^4}{40} + \mathcal{O}(n^{-3/2}).$$

Also,

$$\begin{aligned}&S^{(2)}(\tilde{z}) = n^3/6 - 3x^2n^2/20 + \mathcal{O}(n^{3/2}),\\ &S^{(3)}(\tilde{z}) = \mathcal{O}(n^{7/2}),\\ &S^{(4)}(\tilde{z}) = -n^5/20 + \mathcal{O}(n^4),\\ &S^{(l)}(\tilde{z}) = \mathcal{O}(n^{l+1}), \quad l \geq 5.\end{aligned}$$

We can now compute (5), for instance by using the classical trick of setting

$$S^{(2)}(\tilde{z})(i\tau)^2/2! + \sum_{l=3}^\infty S^{(l)}(\tilde{z})(i\tau)^l/l! = -u^2/2, \tag{6}$$

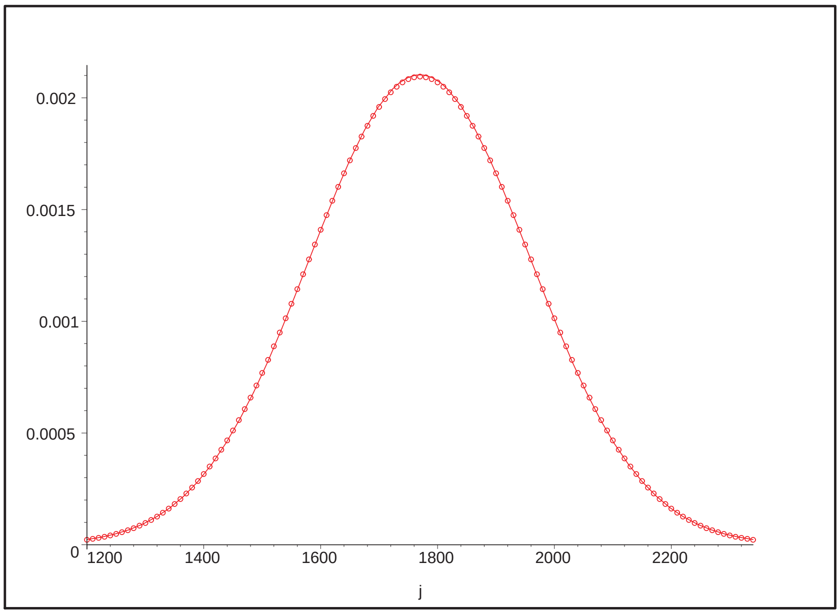

Figure 1: $P_n(j)$ (circle) and the Gaussian asymptotics (line), $n = 60$

computing τ as a truncated series in u, starting as

$$\tau = \frac{u}{n^{3/2}}[\sqrt{6} + \mathcal{O}(1/n)] + \cdots$$

setting $d\tau = \frac{d\tau}{du}du$, expanding w.r.t. n and integrating on $[u = -\infty..\infty]$. (This amounts to the *reversion* of a series.) Finally (5) leads to

$$P_n(j) \sim e^{-x^2/2} \cdot \exp\Big((-9/40 + 9/20x^2 - 3x^4/40)/n + \mathcal{O}(n^{-3/2})\Big)/(2\pi n^3/6)^{1/2}. \tag{7}$$

Note that $S^{(3)}(\tilde{z})$ does not contribute to the $1/n$ correction. Note also that, unlike in the instance of the number of inversions in permutations (see [4]), we have a x^4 term in the first order correction.

To check the effect of the correction, we first give in Figure 1, for $n = 60$, the comparison between $P_n(j)$ and the asymptotics (7), without the $1/n$ term (i.e. the Gaussian contribution). To get a better quality estimation, Figure 2 shows the quotient Q_1 of $P_n(j)$ and the asymptotics (7), without the $1/n$ term. The usual "hat" behaviour is apparent. Figure 3 shows Q_1 and Q_2: the quotient of $P_n(j)$ and the asymptotics (7), with the constant term $-9/(40n)$.
Figure 4 shows Q_2 and Q_3: the quotient of $P_n(j)$ and the asymptotics (7), with the constant term $-9/(40n)$ and the x^2 term $9x^2/(20n)$.
Finally, Figure 5 shows Q_3 and Q_4: the quotient of $P_n(j)$ and the full asymptotics (7). Q_4 gives indeed a good precision on a larger range

For the reader's convenience, here are the quantities that we plot:

$$Q_1 = P_n(j) \Big/ e^{-x^2/2}/(2\pi n^3/6)^{1/2},$$

$$Q_2 = P_n(j) \Big/ e^{-x^2/2} \cdot \exp[(-9/40)/n]/(2\pi n^3/6)^{1/2},$$

$$Q_3 = P_n(j) \Big/ e^{-x^2/2} \cdot \exp[(-9/40 + 9/20x^2)/n]/(2\pi n^3/6)^{1/2},$$

$$Q_4 = P_n(j) \Big/ e^{-x^2/2} \cdot \exp[(-9/40 + 9/20x^2 - 3x^4/40)/n]/(2\pi n^3/6)^{1/2}.$$

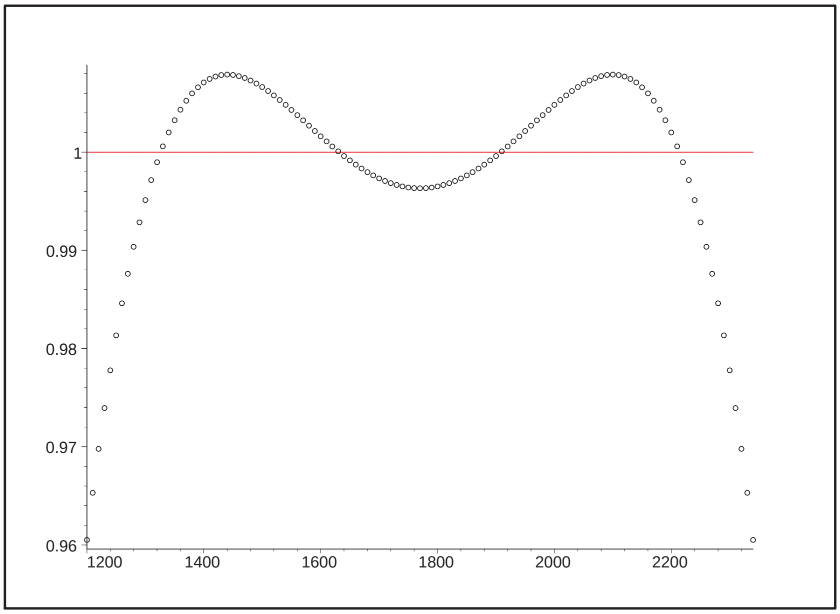

Figure 2: Q_1

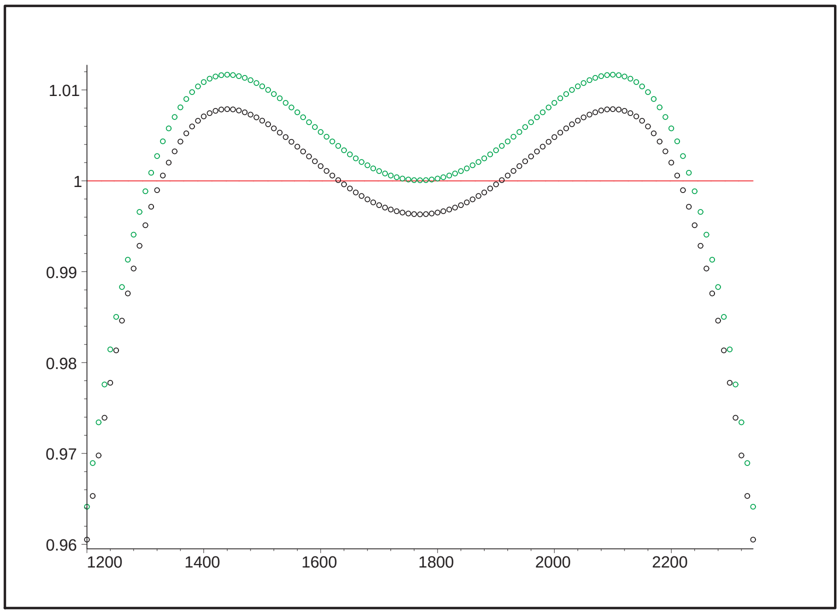

Figure 3: Q_1 (black) and Q_2 (green)

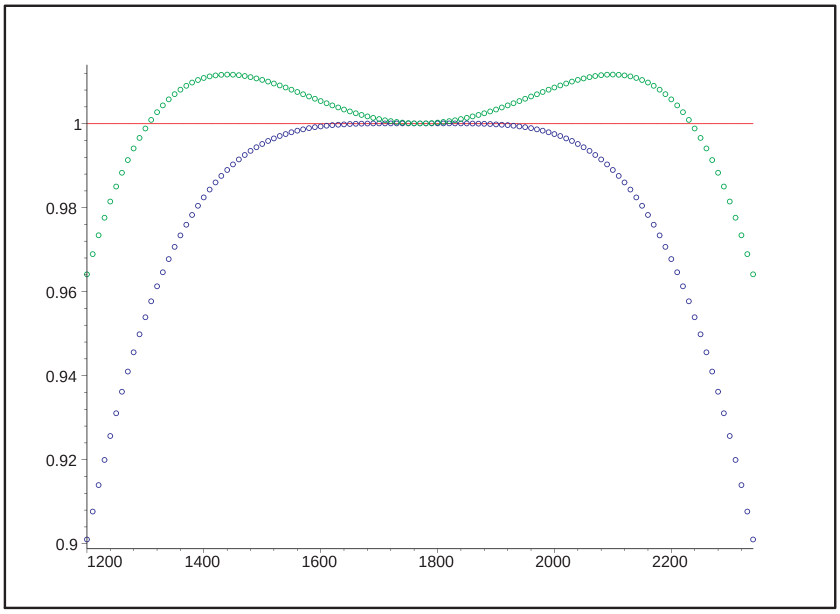

Figure 4: Q_2 (green) and Q_3 (blue)

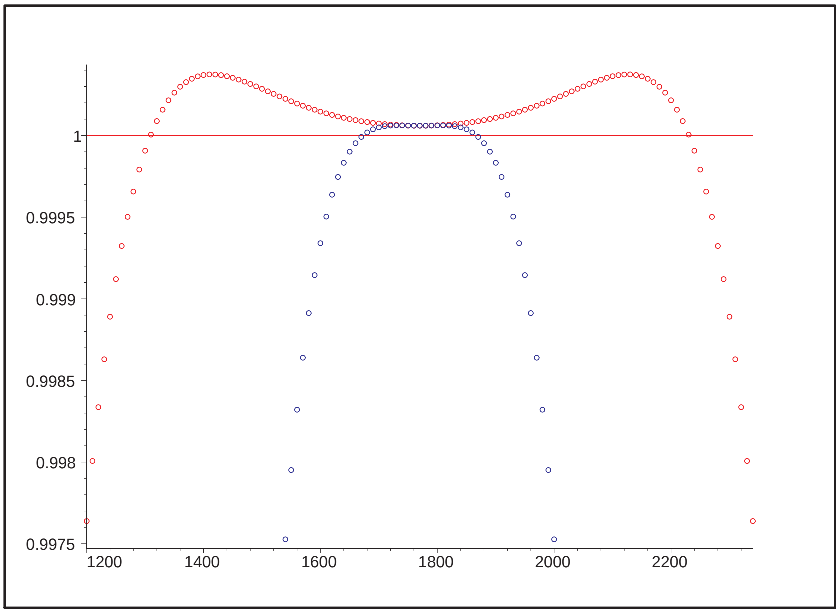

Figure 5: Q_3 (blue) and Q_4 (red)

3 The case $j = n - k$

Now we show that our methods are strong enough to deal with probabilities that are far away from the average; viz. $j = n - k$, for fixed k. Of course, they are very small, but nevertheless we find asymptotic formulæ for them.

We have $z^* = 1$ and, as we will see, ε is again given by a *Puiseux series* of powers of $n^{-1/2}$. But the approach of Section 2 is doomed to failure: all terms of the generalization of (4) contribute to the computation of the coefficients. So we have to turn to another technique. We have

$$P_n(n-k) = [z^{n-k}] \prod_{i=1}^{n-1} \frac{1 - z^{n+i+1}}{1 - z^{i+1}} \Big/ C_n.$$

So we set

$$S := S_1 + S_2 + S_3 + S_4,$$

with

$$\begin{aligned}
S_1 &:= \sum_{i=1}^{n-1} \ln(1 - z^{n+i+1}), \\
S_2 &:= -\sum_{i=1}^{n-1} \ln(1 - z^{i+1}) = -\sum_{i=2}^{n} \ln(1 - z^i), \\
S_3 &:= -(n+1-k)\ln z, \\
S_4 &:= -\ln(C_n).
\end{aligned}$$

3.1 The computation of $\tilde{z}$

Set $\tilde{z} = z^* - \varepsilon = 1 - \varepsilon$. We must have $S'(\tilde{z}) = 0$, with

$$\begin{aligned}
S'(z) &= S_1'(z) + S_2'(z) + S_3'(z), \\
S_1'(z) &= \sum_{i=1}^{n-1} \frac{-(n+i+1)z^{n+i}}{1 - z^{n+i+1}}, \\
S_2'(z) &= \sum_{i=1}^{n-1} \frac{(i+1)z^i}{1 - z^{i+1}}, \\
S_3'(z) &= -\frac{n+1-k}{z}.
\end{aligned}$$

It is clear that we must have either $S_1'(\tilde{z}) = \mathcal{O}(n)$ or $S_2'(\tilde{z}) = \mathcal{O}(n)$. Set

$$\varepsilon = \frac{f(n)}{n},$$

for a function still to be determined. Let us first (roughly) solve

$$S_1'(\tilde{z}) = n.$$

We have

$$S_1'(\tilde{z}) \sim n \frac{\tilde{z}^n}{1 - \tilde{z}} \sim \frac{n^2 e^{-f(n)}}{f(n)}.$$

This leads to the equation

$$\frac{n^2 e^{-f(n)}}{f(n)} = n, \text{ i.e. } e^{f(n)} f(n) = n,$$

which is solved by

$$f(n) = W(n),$$

where W is the Lambert function. But we know that

$$W(n) \sim \ln(n) - \ln\ln(n) + \mathcal{O}\left(\frac{\ln\ln(n)}{\ln(n)}\right).$$

So

$$\varepsilon \sim \frac{\ln(n)}{n}.$$

But then

$$S_2'(\tilde{z}) \sim \frac{1}{(1-\tilde{z})^2} \sim \frac{n^2}{\ln(n)^2} = \Omega(n).$$

So we must turn to the other choice, i.e.

$$S_2'(\tilde{z}) \sim \frac{1}{(1-\tilde{z})^2} \sim \frac{n^2}{f(n)^2}.$$

The equation

$$\frac{n^2}{f(n)^2} = n$$

leads now to

$$f(n) = \sqrt{n},$$

and

$$S_1'(\tilde{z}) \sim \frac{n^2 e^{-\sqrt{n}}}{\sqrt{n}} = n^{3/2} e^{-\sqrt{n}},$$

which is exponentially negligible. This rough analysis leads to the more precise asymptotics

$$\varepsilon = \frac{a_1}{n^{1/2}} + \frac{a_2}{n} + \frac{a_3}{n^{3/2}} + \cdots \tag{8}$$

and we must solve

$$S_2'(\tilde{z}) + S_3'(\tilde{z}) = 0,$$

i.e.

$$\sum_{j=1}^{n-1} \frac{(j+1)\tilde{z}^j}{1-\tilde{z}^{j+1}} - \frac{n+1-k}{\tilde{z}} = 0.$$

We rewrite this equation as

$$\sum_{j=2}^{n} \frac{j\tilde{z}^j}{1-\tilde{z}^j} = M,$$

and will later replace M by $n+1-k$. It is not hard to see that we can solve

$$\sum_{j=2}^{\infty} \frac{j\tilde{z}^j}{1-\tilde{z}^j} = M,$$

instead, since the extra terms only introduce an exponentially small error.

Now we replace $\tilde{z}$ by e^{-t}:

$$g(t) = \sum_{j=2}^{\infty} \frac{je^{-jt}}{1-e^{-jt}},$$

and compute its Mellin transform (for a good reference on Mellin transforms, see Flajolet et al. [2] or Szpankowski [7]):

$$g^*(s) = (\zeta(s-1)-1)\zeta(s)\Gamma(s). \tag{9}$$

The original function can be recovered by a contour integral:

$$g(t) = \frac{1}{2\pi i}\int_{3-i\infty}^{3+i\infty} (\zeta(s-1)-1)\zeta(s)\Gamma(s)t^{-s}ds.$$

This integral can be approximately (in a neighbourhood of $t = 0$) solved by taking residues into account:

$$g(t) = \frac{\pi^2}{6t^2} - \frac{3}{2t} + \frac{13}{24} - \frac{t}{12} + \frac{t^3}{720} - \frac{t^5}{30240} + \mathcal{O}(t^7).$$

Now we solve

$$\frac{\pi^2}{6t^2} - \frac{3}{2t} + \frac{13}{24} - \frac{t}{12} + \frac{t^3}{720} - \frac{t^5}{30240} \sim M,$$

and find

$$t(M) \sim \frac{\pi}{\sqrt{6}}M^{-1/2} - \frac{3}{4}M^{-1} + \frac{(81+13\pi^2)\sqrt{6}}{288\pi}M^{-3/2} - \Big(\frac{13}{32} + \frac{\pi^2}{144}\Big)M^{-2}.$$

To simplify the next expressions, we will set $n = w^2$.

Now, since $1 - \varepsilon = e^{-t}$, we set $M = n + 1 - k$, expand and reorganize:

$$\varepsilon \sim \frac{\pi}{\sqrt{6}}w^{-1} - \Big(\frac{\pi^2}{12} + \frac{3}{4}\Big)w^{-2} + \frac{\sqrt{6}(4\pi^4 + 3\pi^2 + 72\pi^2 k + 243)}{864\pi}w^{-3}. \tag{10}$$

Note that k appears (linearly) only in the coefficient of w^{-3}. The quadratic term in k appears in the coefficient of w^{-5}.

3.2 The computation of $S(\tilde{z})$

First we notice that

$$S_1(z) = \sum_{i=1}^{n-1} \ln(1 - z^{n+i+1}) = \mathcal{O}(e^{-\pi w/\sqrt{6}}w),$$

which is exponentially negligible. Next, instead of using

$$S_2(\tilde{z}) = -\sum_{i=2}^{n} \ln(1 - \tilde{z}^i),$$

we can use

$$\varphi_0(\tilde{z}) := -\sum_{i=2}^{\infty} \ln(1 - \tilde{z}^i) = -\sum_{i=2}^{\infty} \ln(1 - (1-\varepsilon)^i), \tag{11}$$

again with exponentially small error. Now setting again $\tilde{z} = e^{-t}$:

$$\psi_0(t) = \varphi_0(e^{-t}) = -\sum_{i\geq 2} \ln(1 - e^{-ti}) = \sum_{i\geq 2, l\geq 1} \frac{e^{-ilt}}{l},$$

the Mellin transform of which is

$$\psi_0^*(s) = (\zeta(s) - 1)\zeta(s+1)\Gamma(s),$$

and, by residues,

$$\psi_0(t) \sim \frac{\pi^2}{6t} + \frac{3}{2}\ln(t) - \frac{1}{2}\ln(2\pi) - \frac{13t}{24} + \frac{t^2}{24}.$$

Now, with $t = -\ln(1-\varepsilon)$ and (10),

$$S_2(\tilde{z}) \sim \frac{\pi w}{\sqrt{6}} - \frac{5}{4}\ln(2) + \ln(\pi) + \frac{3}{4} - \frac{3}{4}\ln(3) - \frac{3}{2}\ln(w) - \frac{\sqrt{6}(8\pi^2 k + 5\pi^2 + 81)}{96\pi}w^{-1}.$$

Also we easily derive

$$S_3(\tilde{z}) \sim \frac{\pi w}{\sqrt{6}} - \frac{3}{4} - \frac{\sqrt{6}(24\pi^2 k - 37\pi^2 - 81)}{288\pi} w^{-1},$$
$$S_4(\tilde{z}) \sim -\ln(4)w^2 + \frac{\ln(\pi)}{2} + 3\ln(w) + \mathcal{O}(1/w^2),$$

and finally

$$S(\tilde{z}) \sim S_2(\tilde{z}) + S_3(\tilde{z}) + S_4 \sim T_1 + \frac{3}{2}\ln(w) + T_2/w,$$

with

$$T_1 = -2\ln(2)w^2 + \frac{\sqrt{6}\pi w}{3} + \frac{3}{2}\ln(\pi) - \frac{5}{4}\ln(2) - \frac{3}{4}\ln(3), \tag{12}$$
$$T_2 = -\frac{\sqrt{6}(24\pi^2 k - 11\pi^2 + 81)}{144\pi}. \tag{13}$$

3.3 The computation of $S^{(j)}(\tilde{z})$, $j \geq 2$

Set

$$\varphi_j(z) := \varphi_0^{(j)}(z),$$

with φ_0 given by (11). First, we remark that, for any function $\varphi(z)$, with $z = e^{-t}$, $\psi(t) = \varphi(e^{-t})$,

$$\partial_z \varphi(z) = -\partial_t \psi(t) e^{-t},$$

and

$$z\partial_z \varphi(z) = -\partial_t \psi(t).$$

So, if we start with the Mellin transform

$$\psi^*(s) = \int_0^\infty t^{s-1}\psi(t)dt,$$

we have

$$-\int_0^\infty t^{s-1}\partial_t\psi(t)dt = -\psi(t)t^{s-1}\Big|_0^\infty + \int_0^\infty (s-1)t^{s-2}\psi(t)dt = (s-1)\psi^*(s-1).$$

We will check later on that

$$-\psi(t)t^{s-1}\Big|_0^\infty = 0, \tag{14}$$

for the functions we need here.

For instance, if we start from $\varphi_0(z)$ as given by (11), we set

$$G_1(z) = z\varphi_1(z) = z\partial_z\varphi_0(z),$$

and the corresponding Mellin transform is given by

$$\psi_1^*(s) = (s-1)\psi_0^*(s-1) = (\zeta(s-1)-1)\zeta(s)\Gamma(s),$$

which fits with (9).

Similarly the functions

$$G_2(z) = z\partial_z G_1(z) = z\varphi_1(z) + z^2\varphi_2(z),$$
$$G_3(z) = z\partial_z G_2(z) = z\varphi_1(z) + 3z^2\varphi_2(z) + z^3\varphi_3(z),$$
$$G_4(z) = z\partial_z G_3(z) = z\varphi_1(z) + 7z^2\varphi_2(z) + 6z^3\varphi_3(z) + z^4\varphi_4(z),$$

have the corresponding Mellin transforms

$$\begin{aligned}\psi_2^*(s) &= (s-1)\psi_1^*(s-1) = (\zeta(s-2)-1)\zeta(s-1)\Gamma(s),\\ \psi_3^*(s) &= (s-1)\psi_2^*(s-1) = (\zeta(s-3)-1)\zeta(s-2)\Gamma(s),\\ \psi_4^*(s) &= (s-1)\psi_3^*(s-1) = (\zeta(s-4)-1)\zeta(s-3)\Gamma(s),\end{aligned}$$

from which we mechanically derive $\varphi_j(e^{-t})$, $j = 2, 3, 4$. The function φ_5 is not useful here. We must carefully check what precision, in powers of w, we need in order to finally obtain the $1/w^{5/2}$ term in the integration on u (see Section 2). It appears that φ_2 must be computed up to the w^2 term, the other derivatives only need a first term in the expansion. Proceeding as in Section 3.2, including the derivatives of S_3 (S_4 is not used here), we obtain

$$\begin{aligned}S^{(2)}(\tilde{z}) &\sim \frac{2\sqrt{6}w^3}{\pi} + \frac{2(9+2\pi^2)w^2}{\pi^2},\\ S^{(3)}(\tilde{z}) &\sim \frac{36w^4}{\pi^2},\\ S^{(4)}(\tilde{z}) &\sim \frac{144\sqrt{6}w^5}{\pi^3}.\end{aligned}$$

Finally, we can check that the condition (14) is satisfied for all functions we need here.

3.4 The final computation

We compute

$$\tau = \frac{1}{w^{3/2}}\left[\left(\frac{3^{3/4}2^{1/4}\pi^{1/2}}{6} - \frac{2^{3/4}3^{1/4}(9+2\pi^2)}{24\pi^2 w}\right)u - \frac{iu^2}{4w^{1/2}} - \frac{2^{3/4}3^{1/4}u^3}{32\pi^{1/2}w} + \mathcal{O}\left(1/w^{3/2}\right)\right].$$

After integration, we obtain a coefficient

$$\frac{1}{w^{3/2}}\frac{2^{3/4}3^{3/4}}{12}\left(1+\frac{T_3}{w}\right),$$

with

$$T_3 = -\frac{2^{1/2}3^{1/2}(45+8\pi^2)}{48\pi}, \tag{15}$$

and finally,

$$P_n(n-k) \sim e^{T_1}\frac{2^{3/4}3^{3/4}}{12}\left(1+\frac{T_2+T_3}{w}\right) = e^{T_4}\frac{2^{-1/2}\pi^{3/2}}{12}\left(1-\frac{2^{1/2}3^{1/2}(216+13\pi^2+24\pi^2 k)}{144\pi w}\right). \tag{16}$$

with

$$T_4 := -2\ln(2)w^2 + \frac{\sqrt{6}\pi w}{3},$$

The auxiliary quantities are given in (13) and (15).

A plot of the exact value of $P_n(n-k)$, $n = 60$, versus the asymptotics (16) is given in Figure 6. The fit is rather satisfactory, given that we have a linear k contribution only.

3.5 More precision

We have continued the approach of the previous section, in a near-mechanical way, to obtain the w^{-2} and w^{-3} terms, introducing a k^2 and k^3 correction. Let us mention that τ contributes only to the k

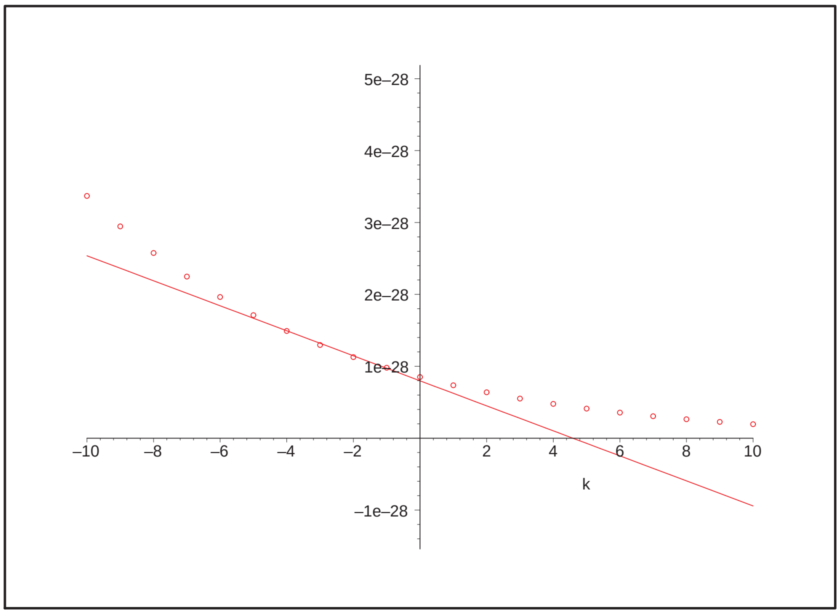

Figure 6: $P_n(n-k), n = 60$ exact (circle), asymptotics (line),

term, the other k contributions come from the $S(\tilde{z})$ expression in the exponent. Omitting the details, this leads to

$$P_n(n-k) \sim e^{T_4} \frac{2^{-1/2}\pi^{3/2}}{12} \Bigg(1 - 2^{1/2}3^{1/2} \frac{216 + 13\pi^2 + 24\pi^2 k}{144\pi w} \\ + \frac{19008\pi^2 + 20736\pi^2 k + 31104 + 217\pi^4 + 624\pi^4 k + 576\pi^4 k^2}{6912\pi^2 w^2} \\ - 3^{1/2}2^{1/2}[-229635 + 11104128\pi^2 + 1907712\pi^4 k + 11197440\pi^2 k \\ +4069\pi^6 + 771984\pi^4 + 1244160\pi^4 k^2 + 15624\pi^6 k + 13824\pi^6 k^3 + 22464\pi^6 k^2]/[2985984\pi^3 w^3] \Bigg).$$

The quality of the approximation is given in Figure 7, with the w^{-2} term (and k^2 contribution) and in Figure 8, with the w^{-2} and w^{-3} terms (and k^3 contribution). The fit is rather good: the curves cover the exact graph above and below.

4 The moments of the Catalan distribution

In [6], a method was described to compute the moments of the distribution given by the Gaussian polynomials. This method can easily be adapted to the instance of the Catalan distribution.

We can write

$$F_n(z) = \frac{g_1(z) \dots g_{2n}(z)}{g_1(z) \dots g_n(z) g_1(z) \dots g_{n+1}(z)},$$

with

$$g_k(z) = \frac{1 - z^k}{k(1-z)},$$

which is the probability generating function of a random variable X_k. The function $\ln g_k(e^{it})$ is the generating function of the cumulants, given by

$$\ln g_k(e^{it}) = \frac{(k-1)it}{2} + \sum_{j \ge 1} \frac{B_{2j}}{2j}(k^{2j} - 1)\frac{(it)^{2j}}{(2j)!}, \qquad |t| < \frac{2\pi}{k}.$$

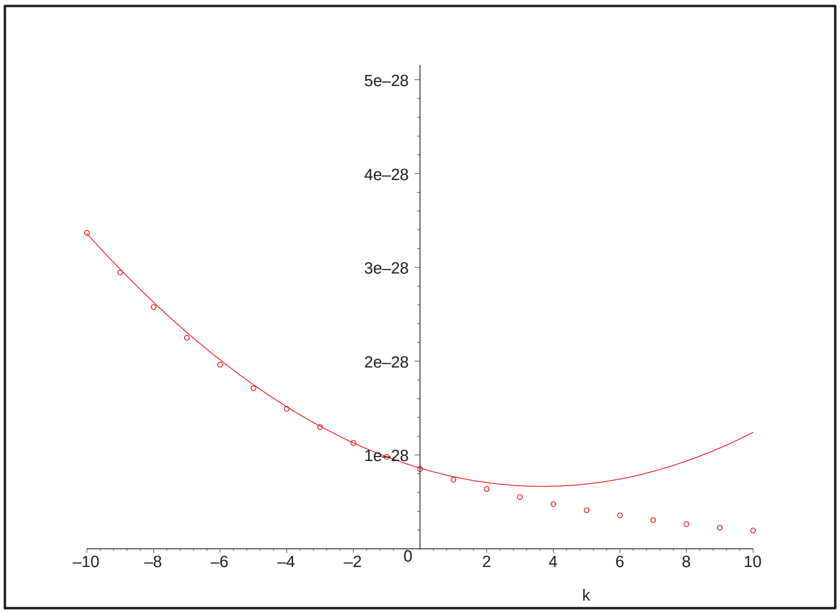

Figure 7: $P_n(n-k), n = 60$ exact (circle), asymptotics (line), with w^{-2} term,

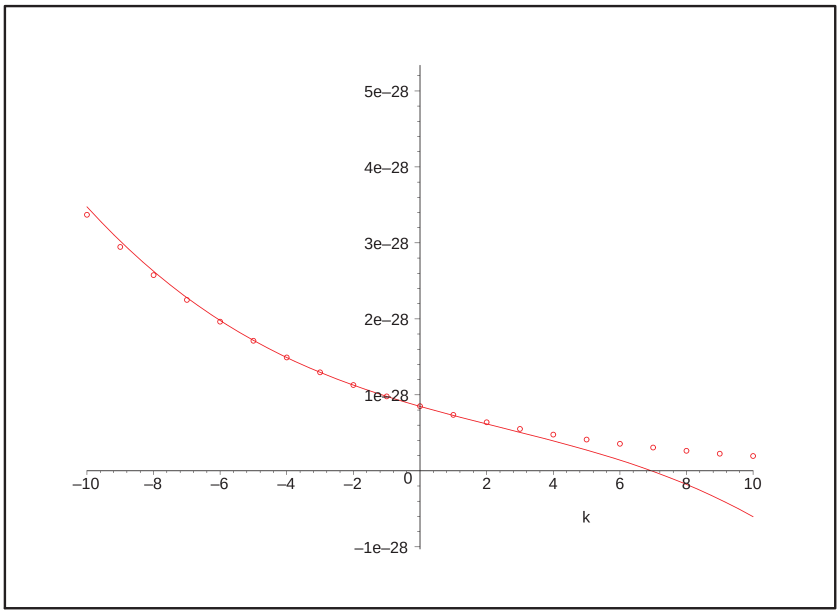

Figure 8: $P_n(n-k), n = 60$ exact (circle), asymptotics (line), with w^{-2} and w^{-3} terms

This gives the cumulants:

$$\kappa_1 = \frac{k-1}{2}, \qquad \kappa_{2r} = \frac{B_{2r}}{2r}(k^{2r}-1), \qquad \kappa_{2r+1} = 0.$$

Since

$$\ln F_n(e^{it}) = \sum_{k=1}^{2n} g_k(e^{it}) - \sum_{k=1}^{n} g_k(e^{it}) - \sum_{k=1}^{n+1} g_k(e^{it}),$$

the cumulants κ_r^n of the Catalan distribution are obtained by summing the κ_r's. Here is a short list:

$$\begin{aligned}
\kappa_1^n &= \frac{n(n-1)}{2},\\
\kappa_2^n &= \frac{n(n-1)(n+1)}{6},\\
\kappa_3^n &= 0,\\
\kappa_4^n &= -\frac{1}{60}n(n-1)(n+1)(3n^2+3n+2),\\
\kappa_5^n &= 0,\\
\kappa_6^n &= \frac{1}{84}n(n-1)(n+1)(6n^4+10n^3+9n^2+5n+2).
\end{aligned}$$

From this, the ordinary moments can be recovered by a standard procedure. Note that κ_1 and κ_2 are the expectation and variance, as predicted by Chen et al. [1].

5 Conclusion

This is another application of the powerful saddle point method in asymptotic enumeration. We treated two instances: around the average, and far away. Other regimes could of course also be treated.

Acknowledgment

Discussions with Peter Grabner are gratefully acknowledged.

References

[1] W. Y. C. Chen, C. J. Wang, and L. X. W. Wang. The limiting distribution of the coefficients of the q-Catalan numbers. Technical report, Nankai University, 2007.

[2] P. Flajolet, X. Gourdon, and P. Dumas. Mellin transforms and asymptotics: Harmonic sums. *Theoretical Computer Science*, 144:3–58, 1995.

[3] P. Flajolet and B. Sedgewick. Analytic combinatorics-symbolic combinatorics: Saddle point asymptotics. Technical Report 2376, INRIA, 1994.

[4] G. Louchard and H. Prodinger. The number of inversions in permutations: A saddle point approach. *Journal of Integer Sequences*, 6:03.2.8, 2003.

[5] A. Odlyzko. *Asymptotic Enumeration*. Eds. R. Graham, M. Götschel, L. Lovász. Elsevier Science, 1995.

[6] H. Prodinger. On the moments of a distribution defined by the Gaussian polynomials. *J. of Statistical Planning and Inference*, 119:237–239, 2004.

[7] W. Szpankowski. *Average Case Analysis of Algorithms on Sequences*. Wiley, New York, 2001.