Can cell cycle distributions be inferred from clone size distributions?

Abstract

Probably not. [1]

In vitro experiments where cells are plated at clonal density may be modelled as a branching process, where each division is independent of all others. It is now routine to accumulate the statistical distribution of the number of progeny (i.e. the clone size) of each initial plated cell. In the theory of super-critical branching processes [2], it is well-known that the limit distribution depends on the lifetime distribution of the particles, i.e. the cell cycle distribution. We therefore seek to infer the cell cycle distribution from the measurable clone size distribution. Concretely, we attempt to invert the relationship between the cell cycle distribution and the clone size distribution, and investigate its stability and (un)suitability as a practical procedure. Along the way we find that not all distributions can arise as the limiting distribution of a super-critical branching process, and we conjecture and prove some properties of them.

Consider an equipotent population of cells which divide independently. This may be modelled as a branching process, defined by specifying the cell cycle distribution be $g(t)$ and the branching outcomes of each division, which we package into a generating function $f(s) = \sum_k p_k s^k$ where p_k is the probability for the cell to divide into k daughters. We will restrict ourselves to biologically feasible processes, in particular we will require all moments of f to exist. Furthermore, we will assume that $p_0 = 0$, as it simplifies the consideration by avoiding total death of the clone; for any given f we can construct a $f^\star$ of equal order such that no death occurs, with the only difference being the existence of terminal clones (see appendixA). We also exclude certain pathological behaviours from consideration, such as the possibility for a clone to reach infinite size in finite time, which excludes in particular the existence of a Dirac delta spike at zero in g. We consider in particular the *super-critical* process, where $f'(1) > 1$ and the expected number of cells diverges exponentially at a rate $\alpha[f, g]$ (the Malthusian parameter of the process), defined by the (only) root of

$$f'(1) \int_0^\infty e^{-\alpha y} g(y) \, dy = 1.$$

In the limit of infinite time, the distribution (normalised to unit mean) of the total number of cells converges to a distribution $H(W)$, whose characteristic function $\phi(u) = \mathbb{E}\left[e^{iuW}\right]$ is the unique solution (amongst distributions with unit mean) of

$$\phi(u) = \int_0^\infty f\left[\phi\left(ue^{-\alpha y}\right)\right] g(y) \, dy. \tag{1}$$

Notice that since the constant α sets a scale for g, we may remove it by scaling g with no change to the limiting behaviour. Therefore without loss of generality we shall set $\alpha = 1$.

We can consider equation 1 as a family of non-linear transform on the space of distributions over the positive reals, indexed by the generating function f. It We note two properties (see appendix C) in particular: it is continuous (in l_1-norm), and injective (for a fixed f, up to scaling of g). In particular, this implies that if we know the branching of cells (which for biological applications we can obtain by experiment), it is in principle possible to undo the transform and infer the cell cycle length from experimental clone size distributions. However, as we will show below, this turn out to be overly optimistic.

In passing we note that equation 1, treated as an iterative procedure, is stable and convergent. However, numerical implementation will have to deal with the inevitable deviation from the manifold of proper characteristic functions. In practice, we find that it is sufficient

to use a cubic spline approximation to ϕ, and clamp the boundary at the origin: $\phi(0) = 1$, $\phi'(0) = i$. With an initial guess corresponding to an exponential distribution for H, this reliably converges to a distribution, although the mean often deviates from unity by a few percent. The convergence is fast enough to be interactive and allow numerical experiments, but only just. We show some examples in appendix B.

To invert the transformation, defining $h(t) = \phi(e^t)$ we obtain

$$h(t) = \int_0^\infty f[h(t-y)]\,g(y)\,dy,$$

which is simply a convolution. Proceeding formally, we may take Fourier transforms and obtain the *Klein inversion formula*:

$$\tilde{g}(\omega) = \frac{\tilde{h}(\omega)}{\widetilde{f \odot h}(\omega)}, \tag{2}$$

where $f \odot h$ is the composition of h followed by f.

Because $\lim_{t \to -\infty} h(t) = \phi(0) = 1$, h is not integrable. Furthermore, since $H(W)$ considered as a distribution on the entire real line almost certainly has discontinuities in its derivatives at the origin, $\phi(u)$ will have algebraic decay for large u; thus $h(t) = O(e^{-ct})$ for $t \to \infty$, for some finite c. Thus the Fourier transform $\tilde{h}$ will only converge and thus be defined on the strip $-c < \Im(\omega) < 0$ (the strip will be as least as wide for $\widetilde{f \odot h}$). Thus for $\tilde{g}$ to be well-defined, we require that the quotient in the inversion formula be analytically extended to include the real line, and give a characteristic function on it; in addition, it must be entirely analytic in the lower half-plane in order for g to be causal, i.e. $g(t) = 0$ for $t < 0$.

Therefore we are led to study what exactly is the image of the transform in equation 1 (equivalently the domain of the Klein formula). We can show than the tail of $H(W)$ is bounded from above by any power-law decay, and conjecture that it is in fact exponential (appendix D).

Directly applying the Klein formula (appendix E), we can reproduce the known results for the exponential (Markovian) process, and the discrete time process. In addition, we can concretely invert the gamma distribution. However, more generally, although formally we can invert distributions of the form $H(W) = p(W)e^{-\lambda W}$ where p is some polynomial, generically they do not yield proper distributions upon inversion. We conjecture that in the space of distributions, at a generic point in the image of the forwards transforms, almost all

directions (i.e. perturbations) lead off the manifold; but nevertheless, there is a countable number of perturbations which remain on it.

Practically, for application to experimental data, the outlook is pessimistic for two reasons. One is that because the transform in equation 1 is continuous, there is not a sharply optimal choice for an inverse that optimises some cost function (likelihood, for instance). In particular, numerical experiments show that changes to g which do not significantly change extreme behaviours tends to have negligible effect on the outcome H. Second is that although the transform is injective, meaning that it is invertible, it depends on having access to the limiting distribution. For interesting problems, such as g having power law tails, a real experiment necessarily will fail to probe the tail structure.

More generically (and realistically), we might have multi-type processes. There we encounter the problem that it would be necessary to have access to the limit distributions starting from all the different cell types, which may present a complete barrier to experiments. It is an open question however, about whether one can approximate a multi-type process with a single-type process, by appropriately choosing f and g.

In conclusion, we find that as a practical procedure, the Klein formula (equation 2) is not practical for biological applications. Nevertheless, it is possible to use it to find some novel pairs of limiting distributions and cell cycle lengths, which would otherwise necessitate solving a non-linear integral equation (equation 1).

Appendix A: Conditioning on survival

In the limit, the probability for extinction is given by the smallest positive root of $q = f(q)$. We can define a new process with

$$f^\star(s) = \frac{f\left[(1-q)\,s + q\right] - q}{1-q}$$

which will yield the a limit distribution $H^\star(W) = (1-q)H(W) + q\delta(W)$. Note in particular that $f^\star$ is a polynomial of the same order as f.

Appendix B: Examples of limiting distributions $H(W)$

First, we consider a simple binary fission $f(s) = s^2$ with Γ-distributed cell cycles

$$g(t) = \frac{\beta^\beta t^{\beta-1} e^{-\beta t}}{\Gamma(\beta)}.$$

The limiting distributions (figure 1) are remarkably close to being Γ-distributed also, but not quite (inset). As we show in appendix E, Γ-distributions are legitimate limiting distributions, but to a slightly different cell cycle distribution.
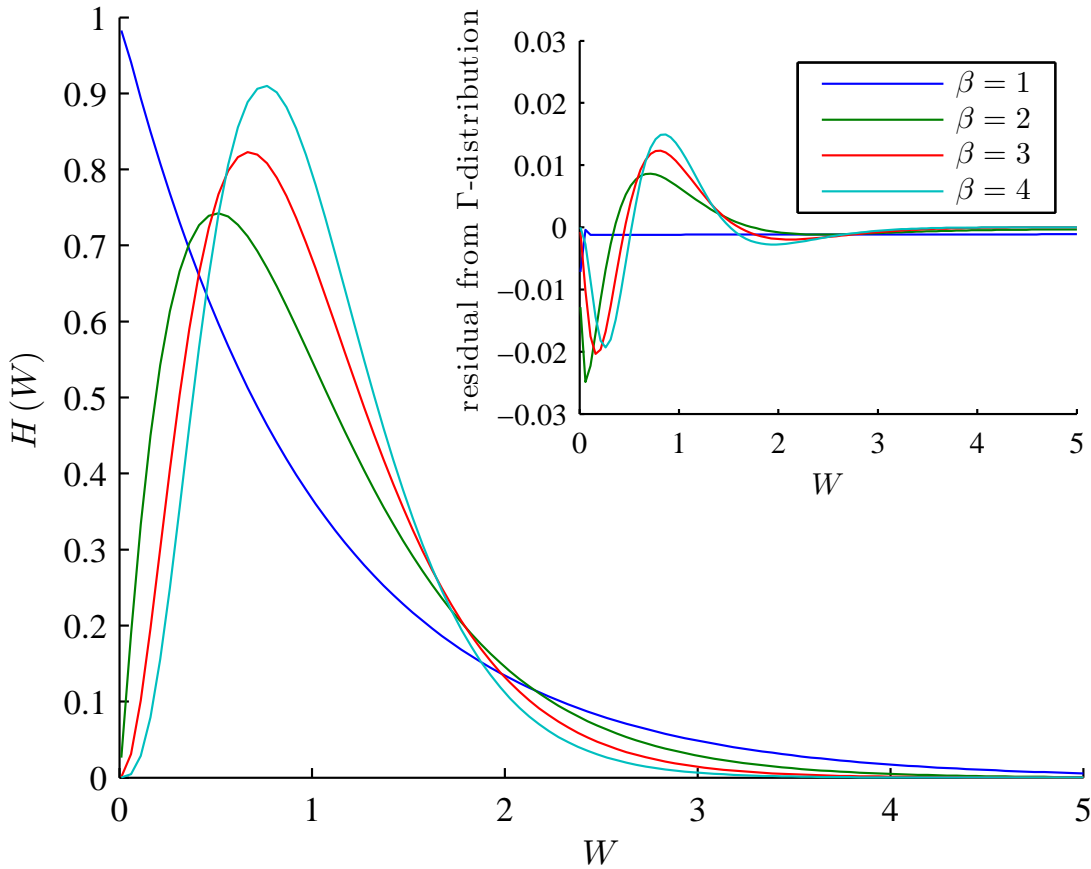

Figure 1. Limit distributions for binary fission with Γ-distributed cell cycle lengths. Inset shows that the limit distributions are very close Γ-distributions, but not exactly.

Since the transform in equation 1 is continuous, the limit distribution will not change qualitatively for moderate changes to the cell cycle distribution. In particular, the biologically relevant example of a mono-modal distribution with an initial refractory period is sufficiently close to give very similar limit distributions. In any case, since the right tail is

always exponential (appendix D), the only qualitative change possible is the behaviour near the origin, i.e. increase the proportion of clones smaller than the average.

First, we may consider a process where the division of each cell can produce a large number of progeny relative to the average. Specifically, consider

$$f(s) = \sum_{k=0}^{\infty} \frac{s^k}{2^k} = \frac{s}{s-2}.$$

For the Markovian process it is possible [2, III.8, theorem 3] to analytically obtain the limit distribution

$$H(W) = \frac{1}{2}\left[-1 + \frac{2e^{-W/4}}{\sqrt{\pi}\sqrt{W}} + \text{erf}\left(\frac{\sqrt{W}}{2}\right)\right],$$

which has a weak divergence at the origin. More intuitively, if we consider the distribution of $\log W$, i.e.

$$J(s) = H\left(e^s\right)e^s = \frac{e^{\frac{s}{2}-\frac{e^s}{4}}}{\sqrt{\pi}} - \frac{1}{2}e^s\text{erfc}\left(\frac{e^{s/2}}{2}\right)$$

$$\sim \frac{1}{\sqrt{\pi}}e^{s/2}, \ s \to -\infty,$$

we see that it still decays exponentially.

Alternatively, we can simply have some cells which divide very slowly. In particular, figure 2 shows the limit distributions for binary fission with the family

$$g(t) = \frac{\text{sinc}\left(\frac{n}{\pi}\right)}{1 + \left(t - \frac{1}{2}\right)^n}\Theta\left(t - \frac{1}{2}\right), \ n \geq 2. \tag{B1}$$

As can be seen in the inset, the corresponding distribution in $\log W$ has a power-law tail for small W, implying a divergence

$$H(W) \approx \frac{C}{W\log^n W}, \ W \to 0.$$

Notice however, that the local minimum in H becomes closer to zero as $n \to \infty$, and the divergence at the origin become more narrow. If we impose an upper-end cut-off to $g(t)$, then the divergence is removed, without significant changes to the rest of the curve.

Appendix C: Injectiveness and continuity of equation (1)

Injectiveness

Assume that two different g_1 and g_2 give the same ϕ upon transform with equation (1). We may assume without loss of generality that their respective Malthusian parameters α
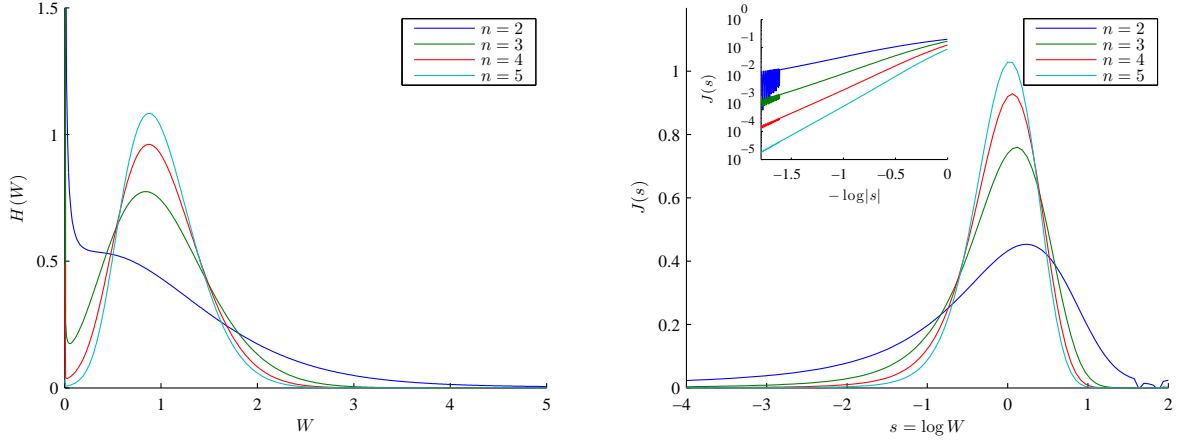
Figure 2. Limit distributions for binary fission power-law tailed cell cycles. Left shows the divergence which occurs at $W \to 0$. Right shows the distribution of $\log W$, which has power-law tails for large negative $\log W$ (inset). The very right tail, and the left tail of the inset shows numerical noise; the distributions are uniformly continuous.

are equal and unity (if they are not we can simply scale the cell cycle distributions). Then we have

$$\int_0^\infty f\left[\phi\left(e^{t-y}\right)\right]\left[g_1\left(y\right) - g_2\left(y\right)\right] dy = 0,$$

for all t. Using the convolution theorem we get

$$\widetilde{f \odot \phi}\left(\omega\right)\left[\tilde{g}_1\left(\omega\right) - \tilde{g}_2\left(\omega\right)\right] = 0$$

for $\Im\left\{\omega\right\} > 0$, obtaining that $\tilde{g}_1$ and $\tilde{g}_2$ are analytic, regular and equal on an open strip. Thus as characteristic functions they are equal, and the original functions g_1 and g_2 are equal as distributions.

Continuity

Having established injectiveness, and since Fourier transforms are continuous, it is sufficient to establish the continuity of equation (2). Letting $h(t) = h_0(t) + \delta h(t)$, we get $f\left[h(t)\right] = f\left[h_0(t)\right] + \delta h\left(t\right) f'\left[h_0\left(t\right)\right]$. Since $0 \le h_0(t) \le 1$, $f'(s)$ is bounded within $0 \le s \le 1$ so define $A = \sup_t f'\left[h_0\left(t\right)\right]$. Taking norms of equation (2), we can straightforwardly com-

pute:

$$
\begin{aligned}
\left\| \widetilde{\delta g} \right\| &= \frac{\left\| \widetilde{\tilde{h}_0} \right\| + \left\| \widetilde{\delta h} \right\|}{\left\| \widetilde{f \odot h_0} \right\| + \left\| \widetilde{\delta h \, f'\left[h_0\right]} \right\|} - \left\| \tilde{g}_0 \right\| \\
&= \left\| \widetilde{\delta h} \right\| \left\| \widetilde{f \odot h_0} \right\| - \left\| \tilde{h}_0 \right\| \left\| \widetilde{\delta h \, f'\left[h_0\right]} \right\| + O\left[\left\| \widetilde{\delta^2 h} \right\| \right] \\
&\leq \left\| \widetilde{\delta h} \right\| \left\{ \left\| \widetilde{f \odot h_0} \right\| + A \left\| \tilde{h}_0 \right\| \right\} + O\left[\left\| \widetilde{\delta^2 h} \right\| \right].
\end{aligned}
$$

Appendix D: Of moments and tails

Differentiating equation 1 we get (via Faà di Bruno's formula):

$$
\phi^{(n)}(u) = \int_0^\infty \left\{ \sum_{\mathbf{k}} \frac{n!}{k_1! \cdots k_n!} f^{(k)} \left[\phi\left(ue^{-y}\right) \right] \left[\frac{\phi^{(1)}\left(ue^{-y}\right) e^{-y}}{1!} \right]^{k_1} \cdots \left[\frac{\phi^{(n)}\left(ue^{-y}\right) e^{-ny}}{n!} \right]^{k_n} \right\} g(y)\, dy
$$

where the summation runs over all partitions of n such that $k_1 + 2k_2 + \cdots + nk_n = n$ and $k = k_1 + \cdots + k_n$. Recognising these as moments (up to factors of i which may be removed without loss of generality by picking the right boundary conditions for ϕ) gives (using the fact that $M_0 = \phi(0) = 1$):

$$
M_n = \sum_{\mathbf{k}} \frac{n!}{k_1! \cdots k_n!} f^{(k)}(1) \left(\frac{M_1}{1!} \right)^{k_1} \cdots \left(\frac{M_n}{n!} \right)^{k_n} \int_0^\infty e^{-ny} g(y)\, dy. \tag{D1}
$$

Since we assumed that all moments of f exist, $f^{(k)}(1)$ will also exist. Noticing that there is only one partition with $k_n = 1$, we can move that term to the left hand side and we get for $n \geq 2$

$$
M_n \left[1 - \int_0^\infty f'(1) e^{-ny} g(y)\, dy \right] = C\left[M_1, \ldots, M_{n-1} \right],
$$

where the right hand side only depends on the lower moments, and in particular is finite if all the lower moments are finite. Then

$$
1 = \int_0^\infty f'(1) e^{-y} g(y)\, dy < \int_0^\infty f'(1) e^{-ny} g(y)\, dy
$$

and since $M_1 = 1$ is finite, by induction all moments exist. Thus, applying Chebychev's inequality we get $H(W) \in o\left(W^{-\alpha}\right)$ for all α. We conjecture that it is actually exponentially bounded, and show this below for binary fission.

We now outline a plausible argument that the tails are also bounded from below by an exponential. Although we do not have proofs, it seems reasonable that some morally similar

8

statements must be true. First, it seems plausible that if two non-negative variables W_1 and W_2 with unity mean, have moments obeying

$$\mathbb{E}\left[W_1^n\right] \geq \mathbb{E}\left[W_2^n\right]$$

for all n, then there exists a w_0 such that for all $w > w_0$

$$\mathbb{P}\left[W_1 > w\right] > \mathbb{P}\left[W_2 > w\right].$$

Second, if the moments $\log \mathbb{E}\left[W^n\right] \in n \log n + \Theta(n)$ for large n then the tail $\mathbb{P}\left[W > w\right]$ is exponential, in some sense. In particular, previous work [3] suggests that if the limits

$$\eta_n = \frac{n\mathbb{E}\left[W^{n-1}\right]}{\mathbb{E}\left[W^n\right]} \to \eta$$

and

$$A_n = \frac{\eta_n^n \mathbb{E}\left[W^n\right]}{n!} \to A$$

exist, then

$$\lim_{w \to \infty} e^{\eta w} \mathbb{P}\left[W > w\right] = A.$$

Finally, we consider equation D1 and drop all terms higher than quadratic

$$\frac{M_n}{n!} \geq \sum_{k=1}^{n-1} \frac{M_k M_{n-k}}{k!(n-k)!} \frac{f''(1) \int_0^\infty e^{-ny} g(y)\, dy}{1 - f'(1) \int_0^\infty e^{-ny} g(y)\, dy} \tag{D2}$$

which follows as each term in equation D1 is positive. We can think of this as a new branching process for which that recurrence relation is exact, which corresponds to a binary fission process with possible death. That process has moments $M_n^\star = n!\xi(n)$ with

$$\xi(n) = \sum_{k=1}^{n-1} \xi(k)\xi(n-k) \frac{f''(1) \int_0^\infty e^{-ny} g(y)\, dy}{1 - f'(1) \int_0^\infty e^{-ny} g(y)\, dy}.$$

Asymptotically, $\int_0^\infty e^{-ny} g(y)\, dy \in \Omega\left(e^{-cn}\right)$ for some c, thus ξ is bounded from below by an exponential. Thus, we conclude that the process with moments $M_n^\star$ has an exponential tail.

In the case of binary fission, we can improve the upper bound slightly. In particular, the equality in equation (D2) applies. An upper bound on the moments is then given by

$$\frac{K_2 M_n}{n!} \leq \sum_{k=1}^{n-1} \frac{K_2 M_k}{k!} \frac{K_2 M_{n-k}}{(n-k)!}.$$

9

with equality for $n \leq 2$, where

$$K_2 = \frac{2 \int_0^\infty e^{-2y} g(y) \, dy}{1 - 2 \int_0^\infty e^{-2y} g(y) \, dy}.$$

We obtain an upper bound in terms of the Catalan numbers

$$\frac{K_2 M_n}{n!} \leq \frac{K_2^n (2n-2)!}{(n-1)! n!}$$

Then applying the moments bound[4] we have

$$\mathbb{P}[W > w] = \inf_n \frac{M_n}{w^n} \leq \inf_n \frac{K_2^{n-1}(2n-2)!}{(n-1)! w^n},$$

which for large w gives

$$\mathbb{P}[W > w] = O\left(\frac{1}{w}\right) e^{-\frac{w}{4K_2}}.$$

Finally, we note that removing the assumption that all moments of f exist will remove the upper bound but not the exponential lower bound. Furthermore, if the k'th moment of f is infinite, then the k'th moment of H will be as well, suggesting a tail $H(W) \approx W^{-k-1}$; indeed, the moment will diverge at finite times.

Appendix E: Applications of the Klein inversion formula

Delta distribution

If $H(W) = \delta(W-1)$ then $\phi(u) = e^{iu}$ and $h(t) = e^{ie^t}$, thus $\tilde{h}(\omega) = e^{\pi \omega/2} \Gamma(-i\omega)$, defined for $0 < \Im(\omega) < 1$. For

$$f(s) = \sum_{k=1}^\infty p_k s^k$$

(notice that we do not have a constant term),

$$\widetilde{f \odot h}(\omega) = e^{\pi \omega/2} \Gamma(-i\omega) \sum_{k=1}^\infty p_k k^{i\omega}$$

defined on the same strip. The quotient is then well defined:

$$\tilde{g}(\omega) = \frac{1}{\sum_{k=1}^\infty p_k k^{i\omega}}.$$

If only one p_k is non-zero (and therefore equal to one), we would have a well-defined distribution

$$g(t) = \delta(t - \ln k),$$

10

which reproduces the trivial result that a discrete time Markovian process with fixed outcomes in each division necessarily leads to a Dirac distribution as the limit. For a more generic f it does not lead to a probability distribution.

Exponential tailed distributions

As conjectured, the tail of H is always exponential. With the integral above we can consider the very general class

$$H(W) = \sum_k a_k \frac{\lambda^{k+1}}{\Gamma(k+1)} W^k e^{-\lambda W}$$

where normalisation requires $\sum_k a_k = 1$ and unity mean requires $\sum_k a_k(k+1)/\lambda = 1$, though as we shall see the latter will be automatically enforced. The characteristic function is very helpfully just a sum

$$\phi(u) = \sum_k a_k \left(1 - \frac{iu}{\lambda}\right)^{-k-1}.$$

Importantly, $f \odot \phi$ has the same structure, composed of a linear combination of $(1 - iu/\lambda)^{-n}$. Thus we only have to consider a single integral (convergent on $-k-1 < \Im\{\omega\} < 0$):

$$\int_{-\infty}^{\infty} e^{i\omega t} \left(1 - \frac{ie^t}{\lambda}\right)^{-k-1} dt$$

$$= \lambda^{i\omega} \int_0^{\infty} z^{i\omega-1}(1 - iz)^{-k-1} dz$$

$$= \lambda^{i\omega} e^{-\pi\omega/2} \frac{\Gamma(1 + k - i\omega)\,\Gamma(i\omega)}{\Gamma(1+k)}$$

$$= \lambda^{i\omega} e^{-\pi\omega/2} \Gamma(1 - i\omega)\Gamma(i\omega) \frac{(1 - i\omega)_k}{\Gamma(1+k)}$$

where $(z)_n$ is the Pochhammer symbol. In equation (2), upon taking the quotient all but the last factor above cancel. Define the auxiliary polynomials $P(z) = \sum_k a_k z^k$ and $Q(z) = \frac{1}{z} f[zP(z)]$ and a transform on polynomials defined by being linear and acts on each monomial as

$$\widehat{z^k} \implies \frac{(1 + z)_k}{\Gamma(k+1)}.$$

Then by linearity we immediately get

$$\tilde{g}(\omega) = \frac{\widehat{P}(-i\omega)}{\widehat{Q}(-i\omega)}.$$

11

Notice that the formula above is completely independent of λ; the correct one such that $H(W)$ has mean of unity will be chose. We can find $g(t)$ by applying the Heaviside formula to $\tilde{g}(iz)$ (treating it as an inverse Laplace transform), and generically it will be a sum of exponentials (possibly with complex exponents, but always occurring in conjugate pairs), determined by the locations of zeros of $\hat{Q}$. Note that whilst $\hat{Q}(z)$ cannot have any zeros on the positive real line, generically there may be zeros on the right half plane. Below, we will try and give some simple cases where this can be done and it yields proper cell cycle distributions.

The simplest case is $H(W) = e^{-W}$ which corresponds to $\lambda = 1$, $P(z) = 1$ and $Q(z) = \frac{1}{z}f(z)$. Thus,

$$\tilde{g}(\omega) = \frac{1}{\sum_k p_k (1 - i\omega)_{k-1}/\Gamma(k)}.$$

If only one p_k is non-zero, then

$$g(t) = (k-1)e^{-(k-1)t} \left(e^t - 1\right)^{k-2}.$$

In particular this reproduces the result that for binary fission ($k = 2$) the Markovian process ($g(t) = e^{-t}$) gives an exponential distribution in the limit. Biologically cells can only divide into two, so we can consider $f(z) = (1 - r)z + rz^2$. In that case,

$$g(t) = \frac{1}{r}e^{-t/r},$$

which is again Markovian, and, in hindsight obvious.

Another simple case is when H is a Γ-distribution; in that case, $\lambda = \beta$, $P(z) = z^{\beta-1}$ and $Q(z) = \frac{1}{z}f\left(\beta^\beta z^\beta\right)$. We get

$$\tilde{g}(\omega) = \frac{(1 - i\omega)_{\beta-1}/\Gamma(\beta)}{\sum_k p_k (1 - i\omega)_{k\beta-1}/\Gamma(k\beta)}$$

$$= \left[\sum_k p_k \frac{(\beta - i\omega)_{(k-1)\beta}}{(\beta)_{(k-1)\beta}}\right]^{-1}.$$

If only one p_k is non-zero then the poles are simple and line on the positive imaginary axis, and furthermore $g(t)$ is positive:

$$g(t) = \frac{\Gamma(k\beta)}{\Gamma(\beta)\Gamma\left[(k-1)\beta\right]}e^{-(k\beta-1)t}\left(e^t - 1\right)^{(k-1)\beta-1}.$$

Note that this reproduces, in appropriate limits, the results above for $H(W) = e^{-W}$ and $H(W) = \delta(W - 1)$. Figure 3 shows some examples from this family.
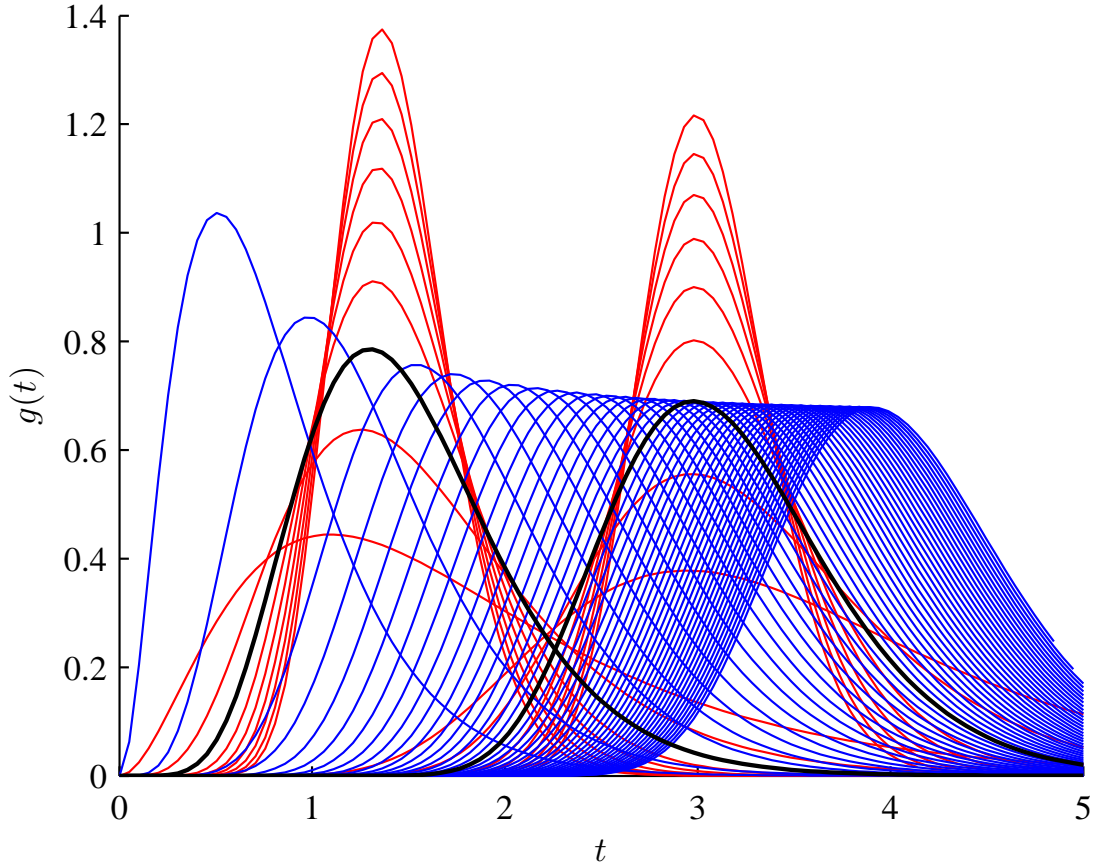
Figure 3. Cell cycle distributions that gives Γ-distributions with shape parameter β as limit distributions, when each cell divides exactly into k at the end of its life. In blue we show the variation as $k = 2, 3, \ldots, 50$ for fixed $\beta = 3$. The two red series show variation of $\beta = 1, \ldots, 9$ for $k = 4$ and $k = 20$. In thick black, we show the intersections $k = 4$, $\beta = 3$ and $k = 20$, $\beta = 3$. As $k \to \infty$ the distributions approach a Dirac delta distribution, with mean $\ln k$.

We note that the above is well-defined for all positive real β, not only positive integers. In addition, for $k = 2$, it produces a family of distributions which when scaled to have unit mean is well-approximated by the Γ-distributions $\beta^\beta e^{-\beta t} t^{\beta-1}/\Gamma(\beta)$. Lastly, for $\beta = 1/(k-1)$, we get an exponential distribution for g, thus solving the Markovian problem for branching with a single k.

More generically, it becomes very difficult to find the conditions such that the Klein formula is well-defined. Consider even the restricted case that $f(z) = (1-r)z + rz^2$, i.e. a

13

generic biologically sensible branching process, starting with a Γ-distribution:

$$\tilde{g}\left(\omega\right) = \left[(1-r) + r\frac{(\beta - i\omega)_\beta}{(\beta)_\beta}\right]^{-1}.$$

Starting with $\beta = 2$, we find

$$g(t) = \frac{6e^{-\frac{1}{2}\left(5 + \sqrt{25 - \frac{24}{r}}\right)t}\left(-1 + e^{\sqrt{25 - \frac{24}{r}}t}\right)}{\sqrt{r(-24 + 25r)}}$$

which is only positive everywhere if $r > 24/25$, in which case all exponents are real. However, for $\beta = 3$

$$g(t) = \frac{60\left[\omega_1\left(e^{t\omega_2} - e^{t\omega_3}\right) + \omega_2\left(e^{t\omega_3} - e^{t\omega_1}\right) + \omega_3\left(e^{t\omega_1} - e^{t\omega_2}\right)\right]}{r\left(\omega_1 - \omega_2\right)\left(\omega_2 - \omega_3\right)\left(\omega_3 - \omega_1\right)}$$

where $\omega_{1,2,3}$ are the three roots to the cubic $60 + 47rz + 12rz^2 + rz^3 = 0$. All three roots have negative real parts for $r \gtrsim 0.106383$ but $g(t)$ is not positive for most of that. Indeed, only for $r > \frac{90\left(270 - \sqrt{3}\right)}{24299} \approx 0.993626$ are all three roots real. Numerically, it seems that only if all three roots are real is $g(t)$ well-defined. By similar arguments, for $\beta = 4$ we find that $r > \frac{840}{841}$; and for $\beta = 5$, $r \gtrsim 0.999906$.

The example above required that the poles of $\tilde{g}$ lay only on the imaginary axis, but this is not necessary. Consider binary fission, $f(z) = z^2$, and a multi-modal $P(z) = r + (1-r)z^3$. The denominator $Q(z)$ is of order 7. The roots are all real for $r \lesssim 9.164 \times 10^{-4}$ and five real roots for $r \lesssim 0.07036$. However, as figure 4 shows, there exist well-behaved solutions for $r \lesssim 0.35$, which can be interestingly multi-modal.

---

[1] M. V. Berry, N. Brunner, S. Popescu, and P. Shukla, Journal of Physics A: Mathematical and Theoretical **44**, 492001 (2011).

[2] K. B. Athreya and P. Ney, *Branching Processes* (Dover Publications, 2004).

[3] J. Abate, G. L. Choudhury, and W. Whitt, Operations Research **44**, 758 (1996).

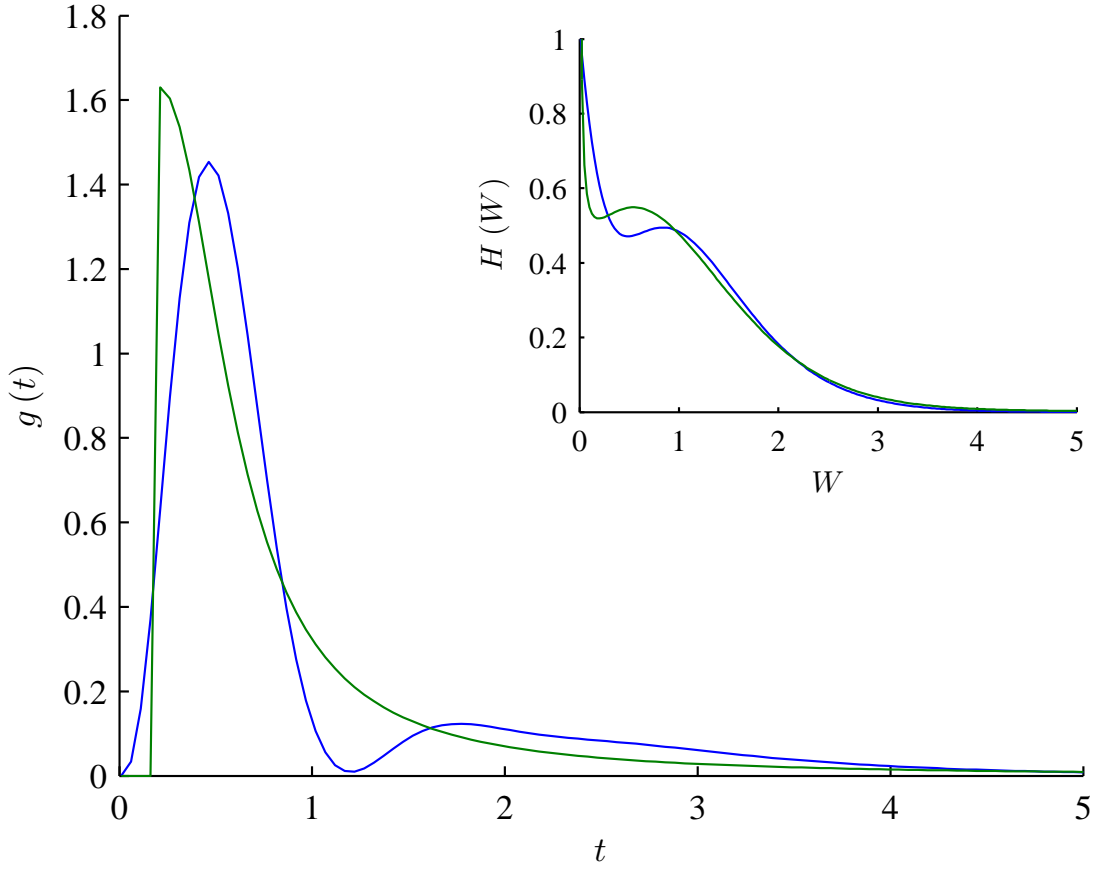[4] T. K. Philips and R. Nelson, The American Statistician **49**, 175 (1995).

Figure 4. Main figure shows the cell cycle distribution which gives the limit distribution in the inset. Blue is for $H(W) = \left[r + (1-r)z^3 W^3 \lambda^4 / \Gamma(4) \right] e^{-\lambda W}$ where λ is such that $\int W \, H(W) \, dW = 1$ and $r = 0.35$. Green is a comparison with the power-law tailed distribution from equation (B1) with $n = 2.1$. For $r \gtrsim 0.35$ the dip in $g(t)$ becomes pronounced enough that it become negative. For $r \to 0$ the distribution $g(t)$ is well-defined, but for $r \to 1$ it is not, even though $r = 1$ it is well-defined.