Transição para o Regime Crítico:

Componentes Gigantes

Quando $p = \frac{1}{N}$, e portanto $\langle k \rangle = p(N-1) \simeq pN = 1$

Aparecem componentes conectados com número de nós da ordem de N, chamados de componentes gigantes. Vamos mostrar essa transição aqui.

Seja N_G o número de nós na componente gigante (CG). Então

$$\mu = 1 - N_G/N = \text{fração de nós fora da CG}$$

$$S = N_G/N = \text{fração de nós na CG} = 1-\mu$$

Se $i \in CG$, ele deve se ligar à outro nó $j \in CG$.

A probabilidade que i NÃO pertença à CG, que deve ser igual à μ, é:

$$\mu = \Big[\underbrace{(1-p)}_{\substack{\text{não há link}\\ \text{entre } i \text{ e } j\text{, mesmo}\\ \text{que } j \in CG.}} + \underbrace{p\mu}_{\substack{\text{existe conexão}\\ \text{entre } i \text{ e } j \text{ mas}\\ j \notin CG}}\Big]^{N-1}$$

deve valer p/ todo j (*)

$$\mu = \left[1 - p(1-\mu)\right]^{N-1} = \left[1 - \frac{\langle k \rangle}{N}(1-\mu)\right]^{N-1}$$

(*) Se existir link entre i e j, então j deve estar fora da CG.

Tirando o log:

$$\log u = (N-1) \log \left[1 - \frac{\langle k \rangle}{N} (1-u) \right]$$

$$\simeq -(N-1) \left[\frac{\langle k \rangle}{N} (1-u) \right]$$

$$\simeq - \langle k \rangle (1-u)$$

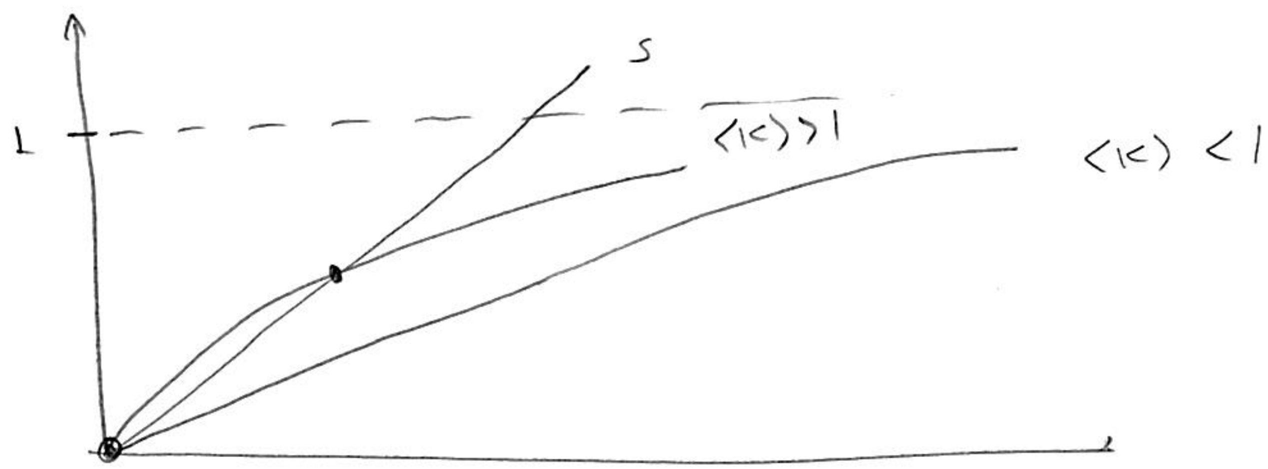
$$u = e^{-\langle k \rangle (1-u)}$$ ou

$$\boxed{1 - S = e^{-S\langle k \rangle}}$$

A solução dessa equação nos dá S, ou $N_G = SN$, p/ cada valor de $\langle k \rangle$, ou $p = \frac{\langle k \rangle}{N}$.

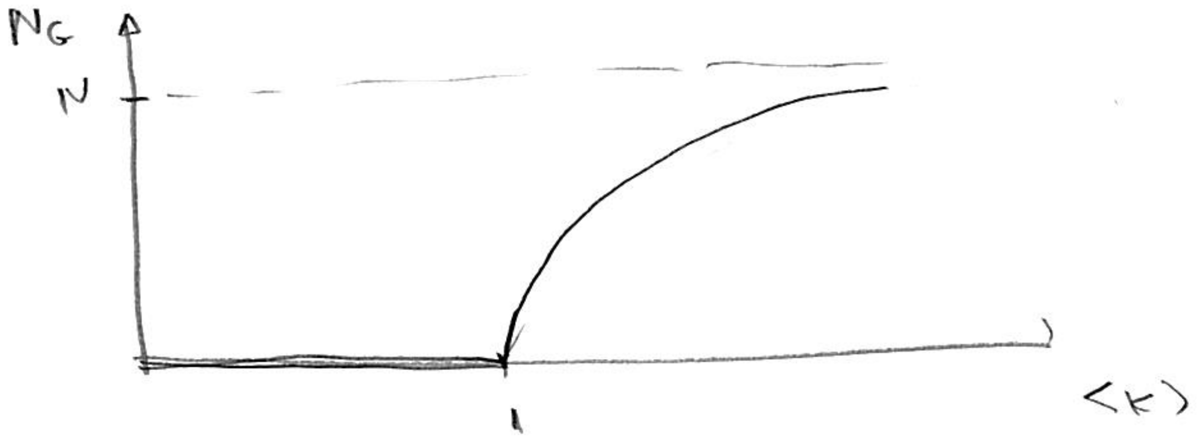
Solução gráfica:

$$S = 1 - e^{-S\langle k \rangle}$$



Se $\langle k \rangle < 1$, $s = 0 \Rightarrow N_G = 0$ e não há formação de componentes grandes.

Se $\langle k \rangle > 1$ existe uma solução com $s > 0$ e $N_G = sN$.



O comportamento perto do ponto de transição pode ser obtido fazendo

$$\langle k \rangle = 1 + \delta k$$
$$s = 0 + A\,\delta k + B\,\delta k^2 + \ldots$$

Substituindo na equação $1 - s = e^{-s\langle k \rangle}$ e expandindo para $\delta k \ll 1$ obtemos $A = 2$ e, portanto, p/ $\langle k \rangle \gtrsim 1$

$$s \simeq 2\,\delta k = 2(\langle k \rangle - 1)$$

TRANSIÇÃO PARA O REGIME CONECTADO

Para $\langle k \rangle > 1$ aparece o componente gigante (CG).

Para qual valor de p teremos um único componente?

A probabilidade de i não pertencer à CG é

$$(1-p)^{N_G} \approx (1-p)^N$$

já que $N_G \approx N$.

O número de nós fora da CG é, portanto,

$$I = N(1-p)^N = N\left(1-\frac{\langle k \rangle}{N}\right)^N$$

$$\approx N e^{-\langle k \rangle} = N e^{-Np}$$

Fazendo $I=1$ temos nossa estimativa

$$1 = N e^{-Np} \Rightarrow N = e^{Np}$$

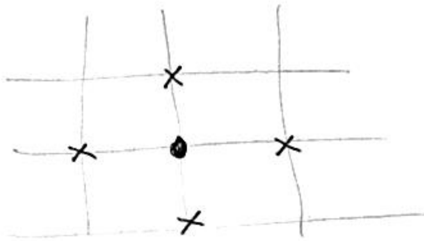
$$\ln N = Np \Rightarrow \boxed{p = \frac{\ln N}{N}}$$

que é ligeiramente maior que $1/N$.

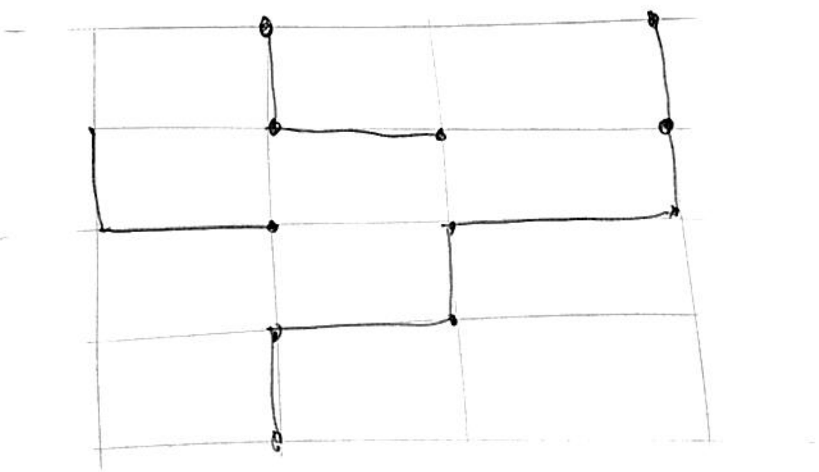
III - Percolação

O processo de evolução da rede aleatória em função da probabilidade de conexão p tem um análogo em teoria de percolação. O problema aparece em várias áreas da Física, como a passagem de água pelo pó de café ou por rochas porosas, ou uma rachadura que se alarga até cobrir toda a extensão de um material.

Vamos considerar uma rede regular 2-D onde cada ponto da rede pode se conectar no máximo com 4 vizinhos:

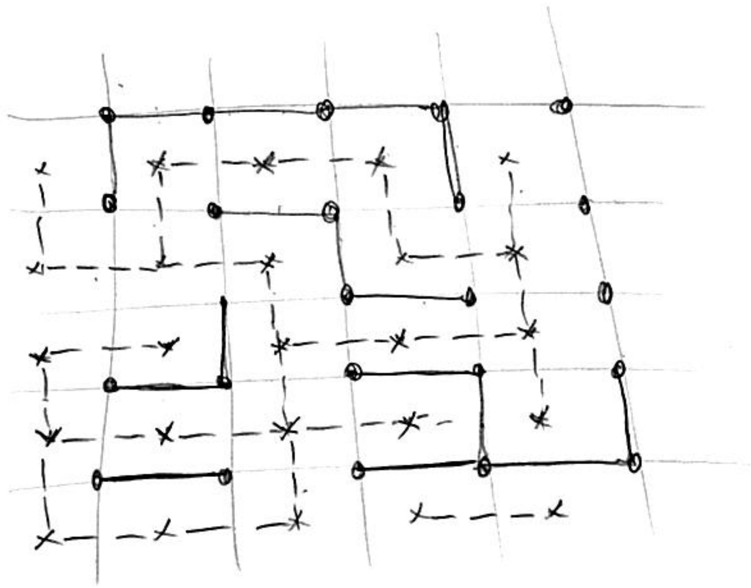


As conexões vão se estabelecer com probabilidade p, como no caso da rede aleatória. O problema que se coloca é: em uma rede muito grande, onde N = número de nós $\to \infty$, qual a probabilidade crítica para o surgimento de um caminho ligando dois lados opostos da rede?



(A) Cálculo exato.

O cálculo de p_c nesse caso pode ser feito exatamente com um truque. Dado um valor de p construímos uma rede como explicado acima:



$N = 25$

Construímos agora uma rede auxiliar onde os nós estão nos centros dos quadrados formados pela rede original.

Os pontos da nova rede serão conectados sempre que não houver uma conexão original bloqueando o link:

rede original ≡ p-rede
rede auxiliar ≡ q-rede

A probabilidade de conexão na q-rede deve ser

$$\boxed{q = 1 - p}$$

$L_p = 16 \rightarrow \langle k \rangle = \frac{16}{25} = 0.64$

$L_{max} = 40 \rightarrow k_{max} = 1.6 \qquad p = \frac{\langle k \rangle}{k_{max}} = 0.4$

$L_q = 25 \qquad \langle k \rangle = 1 \qquad q = \frac{1}{1.6} = 0.625$

As duas redes são do mesmo tipo. Se $p > p_c$ então $q < q_c$. Também, se $q > q_c$, $p < p_c$. Então, quando a p-rede está no ponto crítico, a q-rede também está:

$$q_c = 1 - p_c$$

No entanto, como ambas são redes reticuladas 2-D, tem que ter o mesmo ponto crítico, $q_c = p_c$, o que implica

$$\boxed{p_c = 1/2}$$

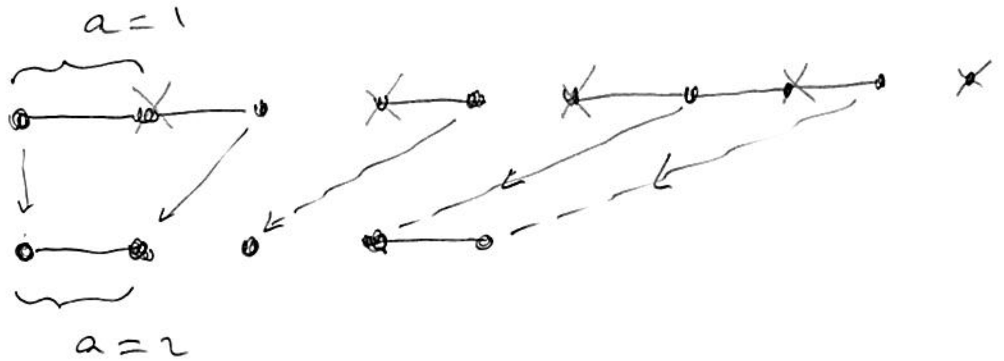
B) Grupo de Renormalização.

Quando $p = p_c$ aparece um caminho conectando dois lados da rede. Se a rede é muito grande o caminho também terá comprimento muito grande. O fenômeno é parecido com a transição líquido-sólido da água a $T_c = 0\,°C$. Um pouco acima de T_c aparecem pequenos aglomerados de gelo, compostos de poucas moléculas de água. Essas moléculas passam a se mover juntas, ficando fortemente correlacionadas. Quando $T \lesssim T_c$ os aglomerados se unem em um mega-aglomerado com 10^{23} moléculas. Dizemos que o comprimento de correlação ficou infinito.

Nesse limite $T = T_c$ ou $p = p_c$ o sistema possui invariância por escala.

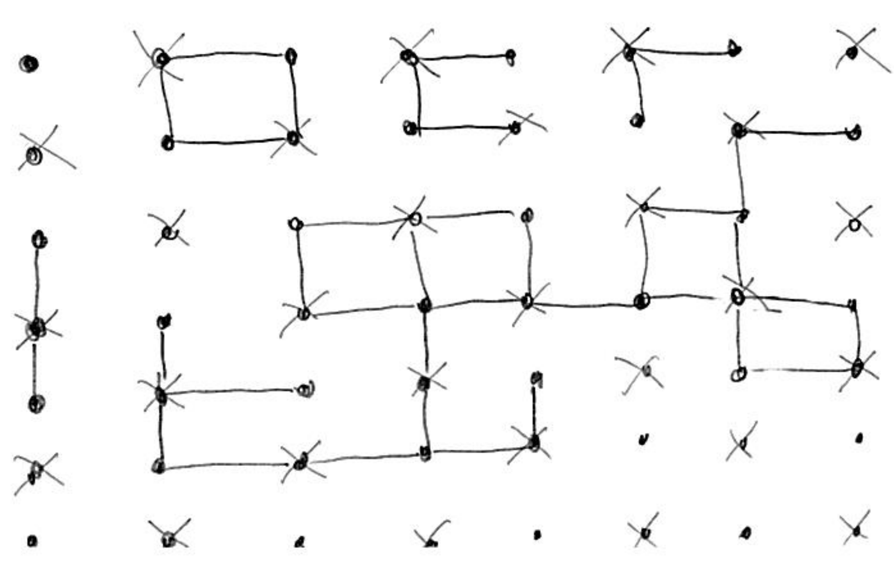
1) Cadeia unidimensional.

Nesse caso $p_c = 1$, caso contrário qualquer não-conexão quebra o caminho. Olhamos agora para a mesma cadeia mas removendo um nó a cada dois



Novos nós serão conectados se estavam conectados antes. A distância entre novos nós vai de $a=1$ para $a=2$. No próximo passo já teremos menos nós desconectados. No entanto, se houvesse um aglomerado gigante, ele sempre existiria em todos os passos do processo.

2) Cadeia bi-dimensional



No ponto crítico a rede deve manter o aglomerado gigante quando fazemos a operação de re-escala. Vamos então remover os sítios marcados com X.

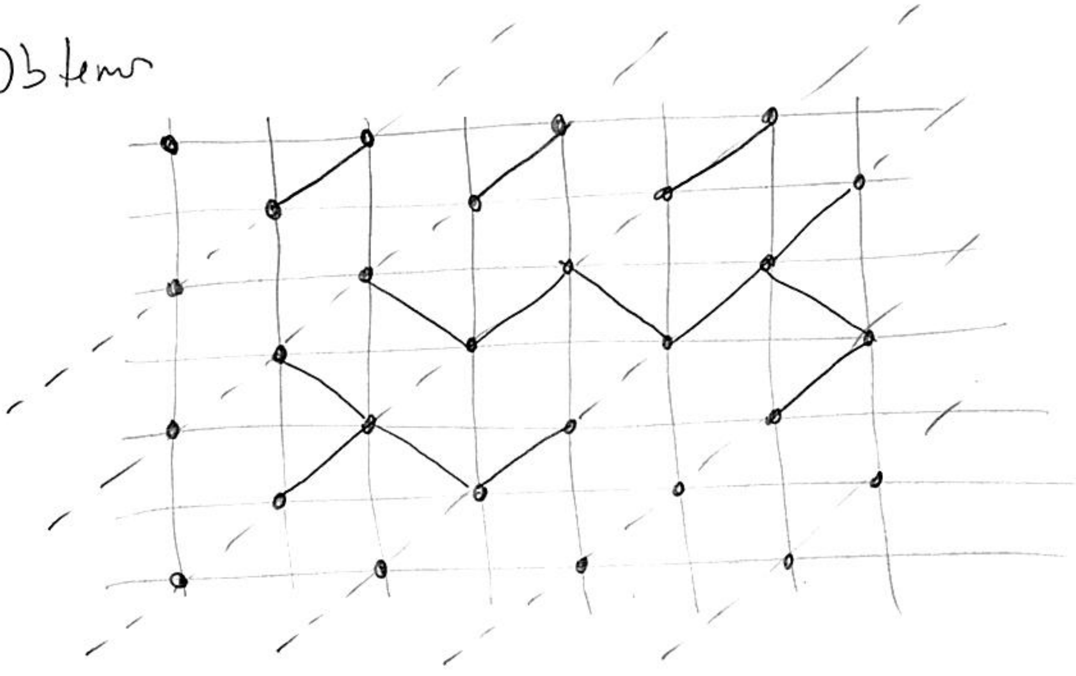
Na nova rede os vizinhos mais próximos de cada sítio se encontram nas diagonais. Dois desses sítios serão conectados nas seguintes situações:

(4 posições possíveis do link aberto)

(2 possibilidades)

 ou 

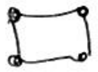
Obtemos



A nova rede roda de 45° e o espaçamento passa de

$$a \longrightarrow a\sqrt{2}$$

Dado que a probabilidade de conexão entre sítios vizinhos é p na rede original, qual o valor da probabilidade de conexão agora, p'?

Probab. de  $= p^4$

Probab. de  $= 4p^3(1-p)$

Prob. de  $= 2p^2(1-p)^2$

$$\Rightarrow \quad p' = p^4 + 4p^3(1-p) + 2p^2(1-p)^2$$

$$= (p^4 - 4p^4 + 2p^4) + (4p^3 - 4p^3) + 2p^2$$

$$= 2p^2 - p^4$$

$$\Rightarrow \quad \boxed{p' = p^2(2-p^2)}$$

No ponto crítico $p' = p$, pois a rede não deve mudar qualitativamente.

R27

As soluções de

$$p = p^2(2-p^2)$$

são $\boxed{p=0}$ e $\boxed{p=1}$ (as triviais) ∴ Escrevendo

$$1 = p(2-p^2)$$

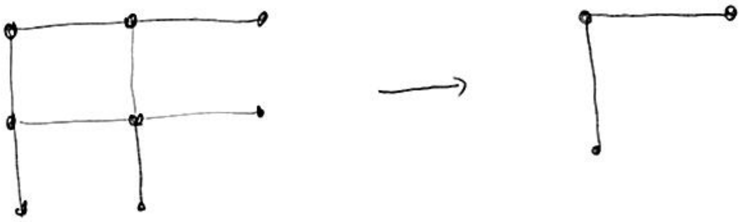
$$p^3 - 2p + 1 = 0 = (p-1)(p^2+p-1)$$

$$\Rightarrow \quad p^2 + p - 1 = 0 \qquad p = -\frac{1}{2} + \frac{1}{2}\sqrt{1+4}$$

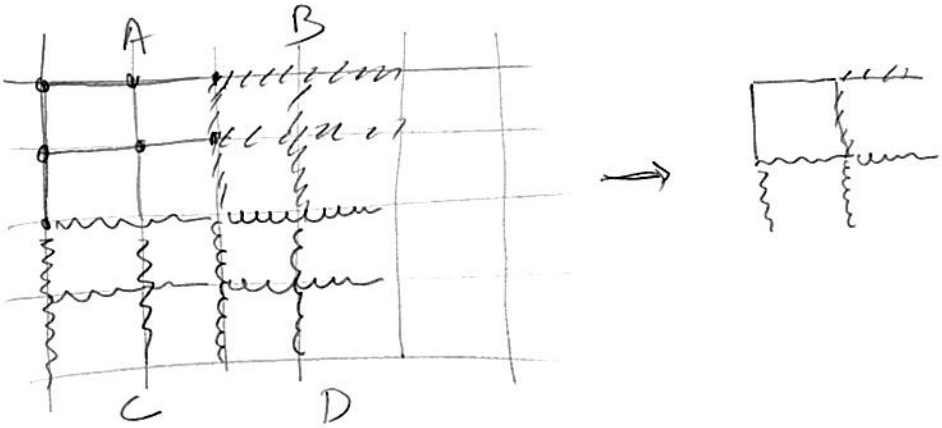
$$p = \frac{\sqrt{5}-1}{2} \simeq 0.62$$

O valor é mais alto do que o exato, mas o procedimento ilustra o método de renormalização, que é a busca do estado invariante por transformações de escala, que é típico de sistemas no ponto de transição de fase.

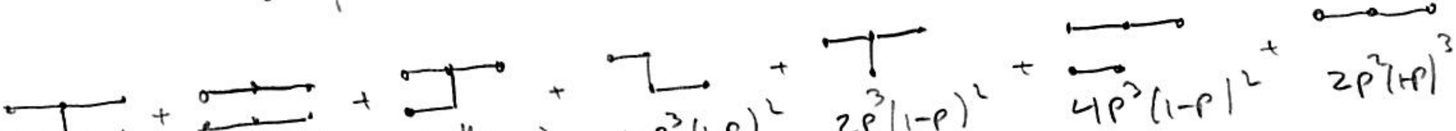
Um esquema mais sofisticado que fornece o resultado exato é mostrado abaixo. Cada grupo de possíveis links à esquerda é levado a um único bloco à direita:



Na rede toda o esquema é



Vamos calcular p' = probabilidade de ter um cluster na direção horizontal, i.e., que cada bloco seja levado em um bloquinho com pelo menos o link horizontal. As possibilidades são



$$p^5 + p^4(1-p) + 4p^4(1-p) + 2p^3(1-p)^2 + 2p^3(1-p)^2 + 4p^3(1-p)^2 + 2p^2(1-p)^3$$

$$p' = 2p^5 - 5p^4 + 2p^3 + 2p^2$$
$$= p^2(2p^3 - 5p^2 + 2p + 2)$$

No ponto crítico $p' = p$ e

$$p\left[2p^4 - 5p^3 + 2p^2 + 2p - 1\right] = 0$$
$$p(p-1)\left[2p^3 - 3p^2 - p + 1\right] = 0$$
$$2p(p-1)(p-1/2)(p^2-p-1) = 0$$

$\rightarrow$ $p = 0$

$p = 1$

$p = 1/2$

$p_\pm = \frac{1 \pm \sqrt{5}}{2}$ $\nearrow p_+ > 1$ $\searrow p_- < 0$

E conseguimos o resultado exato.

Veja que a regra $p' = p^2(2p^3 - 5p^2 + 2p + 2) = f(p)$ é um sistema dinâmico com $f(p^*) = p^*$ para $p^* = 0, 1/2, 1$. Qual a estabilidade desses pontos fixos?

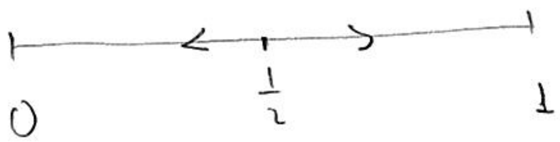
$$\frac{\partial f}{\partial p} = 10p^4 - 20p^3 + 6p^2 + 4p$$

$$\frac{\partial f}{\partial p}(0) = 0$$

$$\frac{\partial f}{\partial p}(1) = 0$$

$$\frac{\partial f}{\partial p}(1/2) = \frac{10}{16} - \frac{20}{8} + \frac{6}{4} + \frac{4}{2} = \frac{10-40+24+32}{16} = \frac{26}{16} = \frac{13}{8} > 0$$

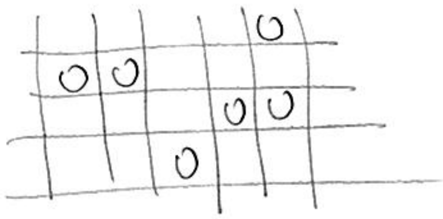
= <u>instável</u>



<u>Exemplo</u> Forest Fire

O = árvore

FOGO PEGA por contato entre vizinhos



- se as árvores são muito densas, o fogo queima a floresta toda (Prob. de ter uma árvore no sítio, P, está acima de P*)

- se as árvores são muito esparsas o fogo queima poucas árvores

MODELO :

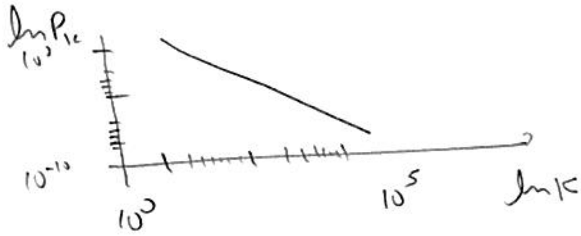
- FOGO APARECE ALEATORIAMENTE com probabilidade f.
- ÁRVORES se REPRODUZEM em sítios VAZIOS VIZINHOS com probabilidade p

Dependendo de f e p o sistema evolui p/ o limiar de percolação. Esse fenômeno é chamado de CRITICALIDADE AUTO-ORGANIZADA.

IV - Redes Livre de Escala

Redes reais NÃO SÃO Aleatórias. Redes como a WWW, Internet, Citações, Proteínas e várias outras NÃO tem distribuição de grau Poisson. A maioria tem

$$P_K \sim K^{-\gamma} \quad \equiv \quad \text{Lei de potência}$$



Tomando o log:

$$\ln P_K \sim -\gamma \ln K$$

Na maioria dos casos $2 < \gamma < 3$.

<u>Definição</u>: redes cuja distribuição de grau é uma lei de potências são chamadas de <u>LIVRE DE ESCALA</u>.

FORMALISMO DISCRETO

$$k = 1, 2, \ldots$$

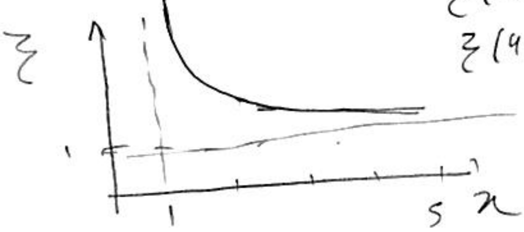
$$P_K = C K^{-\gamma}$$

$$\sum_{K=1}^{\infty} P_K = 1 \quad \Rightarrow \quad 1 = C \sum_{K=1}^{\infty} K^{-\gamma} = C\,\zeta(\gamma)$$

↓ FUNÇÃO ZETA DE Riemann

$$\boxed{P_K = \frac{K^{-\gamma}}{\zeta(\gamma)}}$$

<u>OBS</u>: $\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z}$



$\zeta(2) = \pi^2$

$\zeta(4) = \pi^4/90$

$\zeta(3) \cong 1{,}2020$

FORMALISMO CONTÍNUO

$$P(K) = C K^{-\gamma}$$

$$\int_{K_{min}}^{\infty} P(K)\, dK = 1 \quad \Rightarrow \quad C \int_{K_{min}}^{\infty} K^{-\gamma} dK = \frac{C}{-\gamma+1} K^{-\gamma+1} \Big|_{K_{min}}^{\infty}$$

$$= \frac{C}{\gamma-1} K_{min}^{-\gamma+1} \equiv 1$$

$$C = (\gamma-1) K_{min}^{\gamma-1}$$

$$\boxed{P(K) = (\gamma-1) K_{min}^{\gamma-1} K^{-\gamma}}$$

Veja que $P(K)$ é uma densidade de probabilidade e que

$$\int_{K_1}^{K_2} P(K)\, dK$$ = prob. de um nó ao acaso ter grau entre K_1 e K_2

HUBS : WWW → $N = 10^{12}$, $\langle K \rangle = 4.2$

- Em uma rede aleatória não existem hubs

• Prob. que um nó tenha 100 conexões é

$$P_{100} = e^{-\langle K \rangle} \frac{\langle K \rangle^{100}}{100!} \simeq 10^{-94}$$

• Número de nós com mais de 100 conexões

$$N_{k>100} = 10^{12} \sum_{k=100}^{\infty} \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle} \simeq 10^{-82}$$

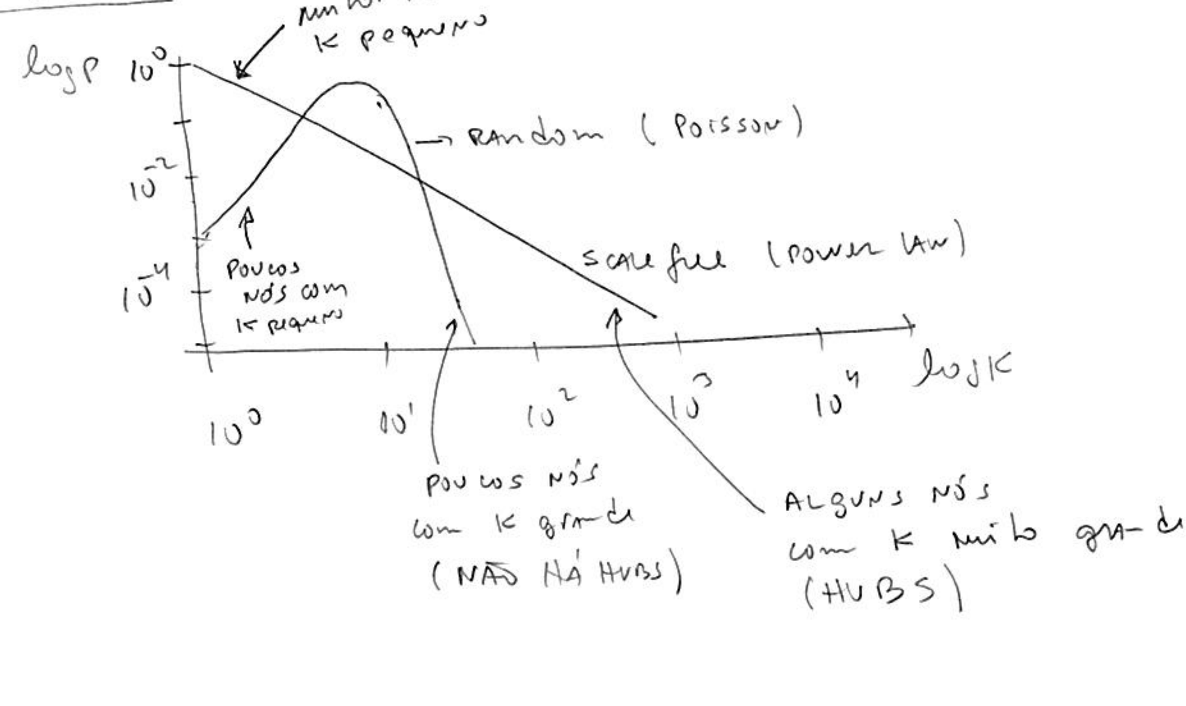
– Usando a lei de potência com $\gamma = 2.1$

$$P_{100} = \frac{100^{-2.1}}{\zeta(2.1)} \simeq 10^{-4}$$

$$N_{k>100} = 10^{12} \sum_{k=100}^{\infty} (\gamma-1)\, k_{min}^{\gamma-1}\, k^{-\gamma}$$

$$\simeq 10^{12} \sum_{k=100}^{\infty} \frac{1}{k^{2.1}} \simeq 10^{12} \frac{1}{100^2} = 10^{8}$$

Graficamente



Uma pergunta importante é: dada uma rede com N nós, qual o grau do nó mais conectado?

MAIOR HUB DA REDE

– Para uma distribuição exponencial

$$P(K) = C e^{-\lambda K}$$

$$1 = \int_{K_{min}}^{\infty} P(K)\, dK = C \frac{1}{(-\lambda)} e^{-\lambda K} \Big|_{K_{min}}^{\infty} = \frac{C}{\lambda} e^{-\lambda K_{min}}$$

$$C = \lambda e^{\lambda K_{min}} \quad \text{e} \quad \boxed{P(K) = \lambda e^{-\lambda (K - K_{min})}}$$

Se a rede tem N nós, estimamos o grau máximo impondo que

$$N \int_{K_{max}}^{\infty} P(K)\, dK = 1$$

ou seja, que exista um nó entre K_{max} e ∞ (prob. de um nó ao acaso ter grau entre K_{max} e ∞ multiplicado pelo nº de nós deve ser 1).

$$N \int_{K_{max}}^{\infty} \lambda e^{\lambda K_{min} - \lambda K}\, dK = N e^{\lambda K_{min}} \lambda \left(-\frac{1}{\lambda}\right) e^{-\lambda K} \Big|_{K_{max}}^{\infty}$$

$$= N e^{-\lambda (K_{max} - K_{min})} \equiv 1$$

$$N = e^{\lambda (K_{max} - K_{min})} \Rightarrow \boxed{K_{max} = K_{min} + \frac{\ln N}{\lambda}}$$

e o crescimento é muito lento com N.

- DISTRIBUIÇÃO POISSON

$$N \sum_{k=k_{max}}^{\infty} P(k) = 1$$

$$N e^{-\langle k\rangle} \sum_{k=k_{max}}^{\infty} \frac{\langle k\rangle^{k}}{k!} \simeq N e^{-\langle k\rangle} \frac{\langle k\rangle^{k_{max}}}{k_{max}!} \equiv 1$$

$$N e^{-\langle k\rangle} = (k_{max})!\ \langle k\rangle^{-k_{max}}$$

tomando o log e usando $\ln n! \simeq n\ln n - n$ obtemos

$$\ln N - \langle k\rangle \simeq \underbrace{k_{max}\ln k_{max}}_{\text{termo dominante}} - k_{max} - k_{max}\ln\langle k\rangle$$

$$\Rightarrow \boxed{k_{max}\ln k_{max} \simeq \ln N}$$

que é ainda mais lento que $k_{max} \sim \ln N$.

- LEI DE POTÊNCIA

$$1 = N\int_{k_{max}}^{\infty} (\gamma-1)\, k_{min}^{\gamma-1}\, k^{-\gamma}\, dk = N(\gamma-1)\, k_{min}^{\gamma-1} \left(\frac{1}{-\gamma+1}\right) k^{-\gamma+1}\Big|_{k_{max}}^{\infty}$$

$$= N\left(\frac{k_{max}}{k_{min}}\right)^{-\gamma+1} \Rightarrow \frac{k_{max}}{k_{min}} = \left(\frac{1}{N}\right)^{\frac{1}{-\gamma+1}}$$

ou

$$\boxed{k_{max} = k_{min}\, N^{\frac{1}{\gamma-1}}}$$

crescimento linear p/ $\gamma = 2$

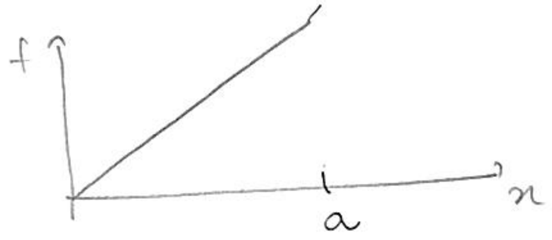
crescimento com $\sqrt{N}$ p/ $\gamma = 3$

O QUE QUER DIZER LIVRE DE ESCALA?

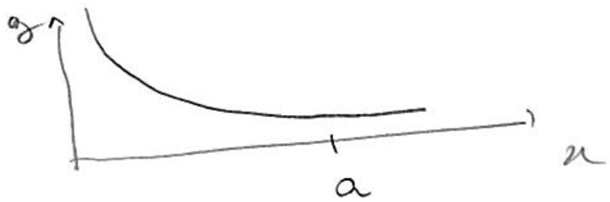
No inglês "SCALE-FREE" seria melhor traduzido por "SEM ESCALA", mas o nome LIVRE DE ESCALA "pegou".

Em primeiro lugar, considere as funções

$f(x) = x$



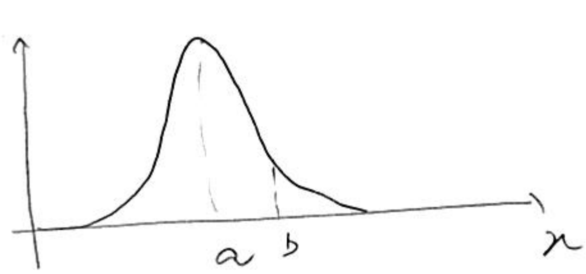
$g(x) = \frac{1}{x}$



Se você desenha essas funções usando $a=1$ ou $a=10$ ou $a=100$, vai ver a mesma figura. Você não consegue identificar o valor de $\underline{a}$ nessas figuras.

A função

$h(x) = e^{-x^2}$



não tem essa propriedade. É claro que $a = 0$ e $b \approx 1$.

$f(x)$ e $g(x)$ não tem escala

Em geral isso ocorre se

$$f(ax) = b\, f(x)$$

que significa que olhar $f(x)$ no ponto ax é igual à $f(x)$, mas com um fator de escala.

Exemplo 1) $f(x) = x$

$$f(ax) = ax = a f(x) \Rightarrow b = a$$

2) $f(x) = \frac{1}{x}$ $\Rightarrow$ $b = \frac{1}{a}$

$$f(ax) = \frac{1}{ax} = \frac{1}{a} f(x)$$

A forma geral da função que satisfaz essa regra é a lei de potência:

$$f(x) = A x^{-\gamma}$$

$$f(ax) = A(ax)^{-\gamma} = a^{-\gamma} f(x)$$

$$\Rightarrow \boxed{b = a^{-\gamma}}$$

Outro aspecto importante são os momentos da distribuição.

$$\langle k^n \rangle = \int_{k_{min}}^{k_{max}} k^n p(k)\, dk = (\gamma-1) k_{min}^{\gamma-1} \int_{k_{min}}^{k_{max}} k^{n-\gamma} dk$$

$$\langle k^n \rangle = (\gamma-1) k_{MIN}^{\gamma-1} \frac{1}{n+1-\gamma} \left[k_{MAX}^{n-\gamma+1} - k_{MIN}^{n-\gamma+1} \right]$$

Para $N \to \infty$, $k_{MAX} \to \infty$ e $\langle k^n \rangle$ diverge se $n-\gamma+1 > 0$

• Se $2 < \gamma < 3$, $n-\gamma+1 > 0$ p/ $n = 2, 3, \ldots$

$$\Rightarrow \quad \langle k \rangle = \frac{(\gamma-1) k_{MIN}}{(\gamma-2)}$$

$\langle k^2 \rangle$, etc $\to \infty$ $\Rightarrow$ NÃO existe variância.

• Se $\gamma > 3$ tanto $\langle k \rangle$ quanto $\langle k^2 \rangle$ convergem, e A rede se comporta como se fosse aleatória.

Na verdade, mesmo p/ $2 < \gamma < 3$ $\langle k^2 \rangle$ só diverge se N é infinito, mas se N é grande $\langle k^2 \rangle$ é muito grande.

• Se $\gamma < 2$ até mesmo $\langle k \rangle$ diverge. Essas redes são anômalas e não devem ocorrer

Pode-se mostrar que a distância média entre dois nós é

$$\langle d \rangle = \begin{cases} \text{const.} & \text{se } \gamma = 2 \\ \ln(\ln N) & \text{se } 2 < \gamma < 3 \\ \dfrac{\ln N}{\ln \ln N} & \text{se } \gamma = 3 \\ \ln N & \text{se } \gamma > 3 \end{cases}$$

COMO GERAR REDES COM DISTRIBUIÇÕES ARBITRÁRIAS

I) Modelo Configuracional

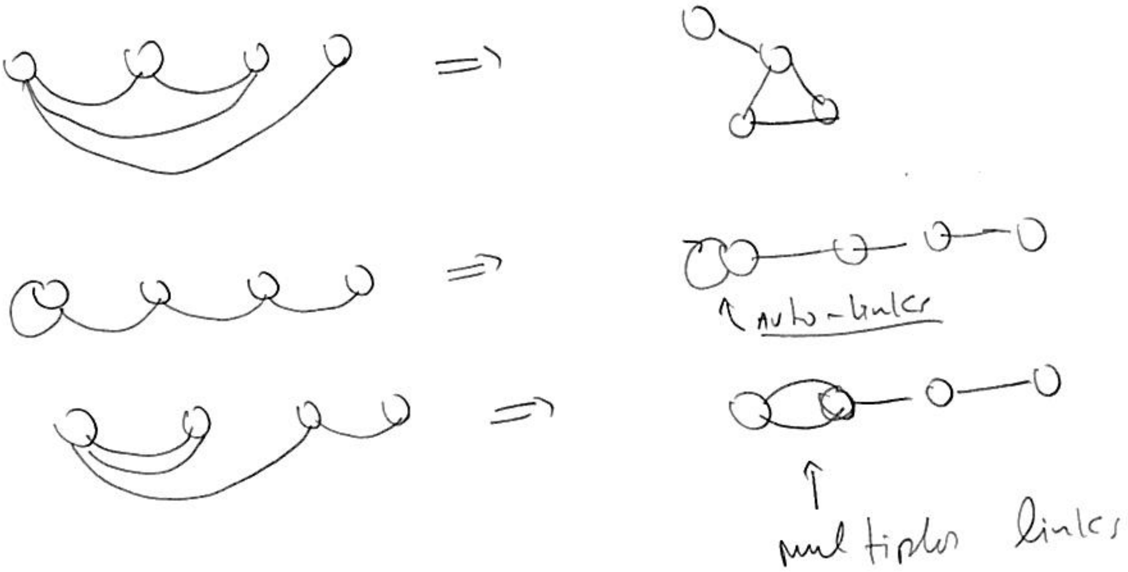
- Começamos com N nós com graus pré-determinados

$k_1=3$ $k_1=2$ $k_3=2$ $k_4=1$

$$\sum k_i = \text{par} = 2L$$

- Sorteamos um par de links e os conectamos

Podemos obter redes diferentes p/ sorteios diferentes:

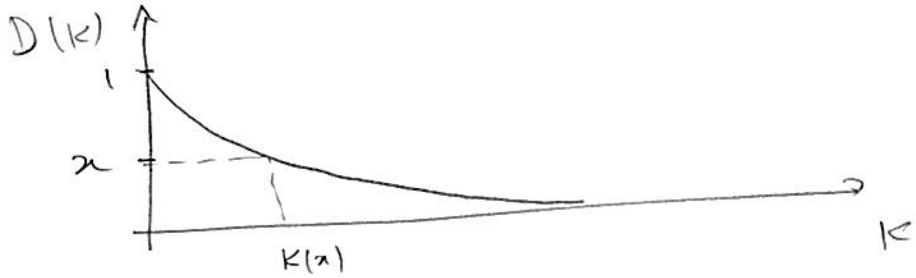


O modelo é útil p/ gerar um conjunto de redes com a mesma distribuição de uma rede original que queiramos estudar.

Como gerar uma sequência de graus que segue uma distribuição $P(k)$?

(1) $$\int_0^\infty P(k)\,dk = 1$$

$$D(k) \equiv \int_k^\infty P(k')\,dk' = \text{prob. de um nó ter grau maior ou igual à } k.$$



(2) Sorteamos agora um número aleatório $x \in [0,1]$ e vemos qual k corresponde à $D(k) = x$.

(3) Fazemos isso p/ vários valores aleatórios de x:

$$\begin{array}{ccccc} x_1, & x_2, & x_3, & \ldots & x_N \\ \downarrow & \downarrow & \downarrow & & \downarrow \\ k_1 & k_2 & k_3 & & k_3 \end{array}$$

Exemplo 1 – Distrib. exponencial

$$P(k) = \lambda e^{-k\lambda} \qquad (k_{min} = 0)$$

$$D(k) = \int_k^\infty P(k')\,dk' = e^{-\lambda k}$$

SORTEANDO x_1 temos $e^{-\lambda k_1} = x_1$

$$\Rightarrow \boxed{k_1 = \frac{1}{\lambda} \ln x_1^{-1}}$$

Exemplos ($\lambda = 1$)

$x_1 = 0.5 \quad k_1 = 0.7$

$x_2 = 0.1 \quad k_2 = 2$

$x_3 = 0.01 \quad k_3 = 5$

Exemplo 2 – Lei de Potências

$$P(k) = (\gamma - 1)\, k^{-\gamma}$$

$$D(k) = k^{-\gamma+1} \qquad (k_{min} = 1)$$

SORTEANDO $x_1 \longrightarrow k_1^{-\gamma+1} = x_1$

$$(\gamma - 1) \ln k_1 = \ln x_1^{-1}$$

$$k_1 = \exp\left\{\frac{1}{\gamma-1} \ln x^{-1}\right\} = x_1^{-(\gamma-1)}$$

$$\boxed{k_1 = \frac{1}{x_1^{\gamma-1}}}$$

$\gamma = 2.1 \quad x_1 = 0.5 \quad k_1 = 2$

$x_2 = 0.1 \quad k_2 = 12$

$x_3 = 0.01 \quad k_3 = 158$

Outro método é

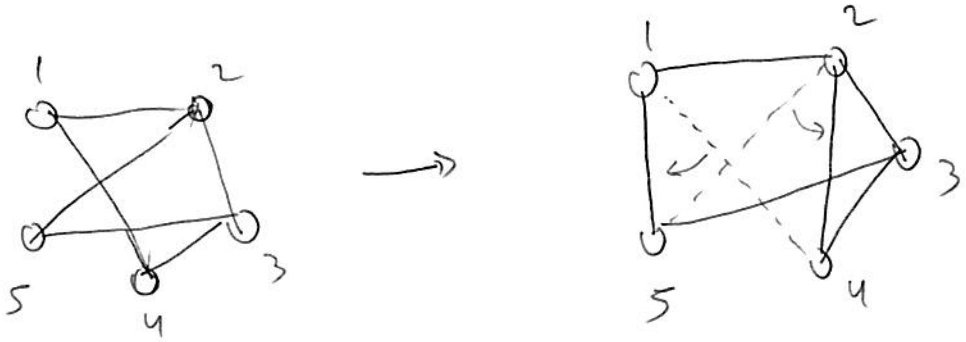
(II) Randomização com Preservação de Grau

Dada uma rede com N nós e graus $k_1, k_2, \dots k_N$, sorteia-se dois links, que são trocados se não geraram auto-links ou multiplos links:

Link 1–4 $\longrightarrow$ Link 1–5

Link 2–5 $\quad\quad\;\;$ Link 2–4

↑ links originais $\quad$ ↑ novos links



V - O MODELO DE BARABÁSI-ALBERT

Como observamos no capítulo anterior, a maioria das redes reais não são aleatórias e várias delas possuem a característica de serem livres de escala, com $P(k) \sim k^{-\gamma}$.

Porque então sistemas tão diferentes como redes de proteínas e a WWW teriam essa mesma propriedade? Barabási e Albert identificaram dois mecanismos comuns à várias redes reais:

- As redes crescem com o tempo, com novos nós sendo acrescentados
- cada novo nó tende a se conectar com outros que já eram muito conectados

Constroe-se então o seguinte modelo, chamado de "<u>CRESCIMENTO COM CONEXÕES PREFERENCIAIS</u>", ou "<u>MODELO BA</u>" (BARABÁSI-ALBERT)

1) A rede começa com m_0 nós, conectados de forma arbitrária. (totalmente conectada, ou sem conexão alguma ou outra distribuição de links).

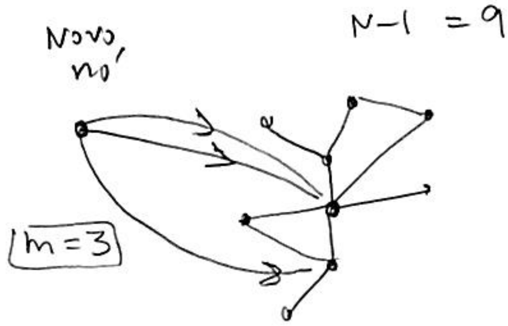
2) A cada passo de tempo um novo nó é adicionado e fará m conexões com os nós já existentes

A probabilidade que cada um desses links se conecte com o nó i é

$$\Pi_i = \frac{K_i}{\sum_{j=1}^{N-1} K_j} = \text{proporcional ao grau do nó } i$$

- $N-1$ = nº de nós da rede sem contar o novo nó que está sendo ligado.
- O modelo é probabilístico. Se $\Pi_1 = 1/5$ e $\Pi_2 = 4/5$ ainda pode ocorrer a ligação com o nó 1.

Note que podem se formar multiplos links entre os mesmos dois nós:



Podemos evitar isso diretamente na construção.

EVOLUÇÃO TEMPORAL DA REDE BA

A cada passo de tempo m novos links são incluídos. O grau K_i do nó i vai mudar (em média) por

$$K_i(t+1) = K_i(t) + m\Pi_i$$

ou

$$K_i(t+1) - K(t) = m\Pi_i$$

Tratando $k_i(t)$ como uma variável contínua

$$\frac{dk_i}{dt} = m\Pi_i = \frac{m k_i}{\sum_{j=1}^{N-1} k_j}$$

$$\sum_{j=1}^{N-1} k_j = 2mt - m$$

↳ os m novos links ainda não colocados

↓ t passos de tempo (incluindo o atual)

↳ m links p/ passo de tempo

↳ cada link contribui p/ o grau dos dois nós que ele liga

$$\frac{dk_i}{dt} = \frac{m k_i}{m(2t-1)} = \frac{k_i}{2t-1} \simeq \frac{k_i}{2t}$$

$$\frac{dk_i}{k_i} = \frac{1}{2}\frac{dt}{t} \Rightarrow \ln\left(\frac{k_i}{k_{i0}}\right) = \frac{1}{2}\ln\left(\frac{t}{t_i}\right)$$

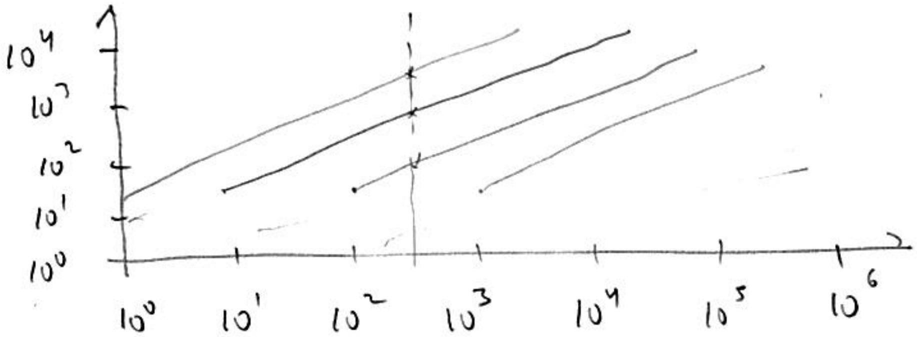
Como cada nó tem m links quando entra na rede

em $t = t_i$,

$$k_i = m\left(\frac{t}{t_i}\right)^{1/2} \text{ ou}$$

$$\boxed{k_i = m\left(\frac{t}{t_i}\right)^{\beta} \quad ; \quad \beta = 1/2}$$

O grau de todos os nó aumenta com $\sqrt{t}$:



Quem entra mais cedo na rede tem sempre mais conexões, isso retarda as conexões dos outros nós, por isso o crescimento é <u>sub-linear</u>:

$$\frac{dk_i}{dt} = \frac{m}{2} \frac{1}{\sqrt{t t_i}}$$

e quanto menor o t_i maior o dk_i/dt.

Distribuição de grau aproximada

Notamos que nós que entraram na rede com

$$t_i < t\left(\frac{m}{k}\right)^{1/\beta} \equiv \bar{t}$$

tem $k_i > k$. De fato, para esses nós,

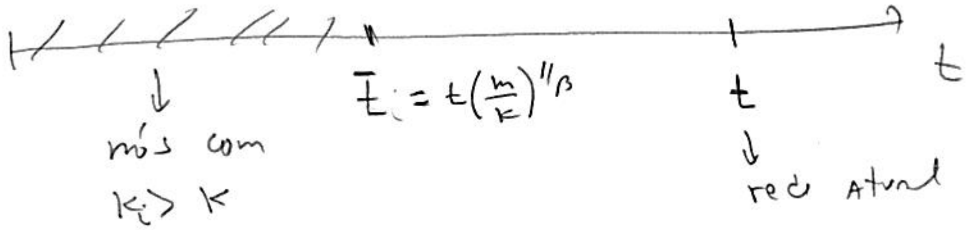
$$k_i = m\left(\frac{t}{t_i}\right)^{\beta} > m\left[\frac{t}{t\left(\frac{m}{k}\right)^{\frac{1}{\beta}}}\right]^{\beta} = m\left[\left(\frac{k}{m}\right)^{\frac{1}{\beta}}\right]^{\beta} = k$$

$\Rightarrow k_i > k$.

e para $t_i = \bar{t}$, $k_i = k$.

Como temos 1 nó a cada passo de tempo, o número total de nós na rede é

$$N = m_0 + t \approx t \quad p/ \ t \text{ grande}$$



de nós com $k_i > k$ é igual à $\bar{t}_i$

$$= t\left(\frac{m}{k}\right)^{1/\beta}$$

Prob. de que um nó da rede tenha grau $> k$ é

$$\frac{\text{N° de nós com grau} > k}{\text{N° total de nós}} = \frac{t\left(\frac{m}{k}\right)^{1/\beta}}{t} = \left(\frac{m}{k}\right)^{\frac{1}{\beta}}$$

$$\Rightarrow \quad P_c(k) = 1 - \left(\frac{m}{k}\right)^{\frac{1}{\beta}}$$

= prob. de um nó ter grau menor ou igual à k

= distribuição cumulativa de grau

$$\boxed{P(k) = \frac{dP_c}{dk} = \frac{m^{1/\beta}}{\beta} k^{-1/\beta - 1}}$$

p/ $\beta = 1/2$

$$\boxed{P(k) = 2m^2 k^{-3}}$$

Essa equação é válida apenas se $k >> 1$ e $m >> 1$.
O erro vem da aproximação contínua que fizemos. Fazemos abaixo a derivação exata. Veja que $P(k) \sim k^{-\gamma}$ e $\gamma = 1 + 1/\beta$, ligando um expoente dinâmico (β) a um topológico (γ).

DISTRIBUIÇÃO DE GRAU EXATA

Sejam $N(k,t)$ = número de nós com grau k no tempo t

$P_k(t)$ = prob. de um nó ter grau k no tempo t
$= N(k,t)/t$

Quando um novo nó é adicionado:

- um link pode ligar a um nó de grau K fazendo $N(k,t) \to N(k,t) - 1$

- um link pode ligar a um nó de grau K−1 fazendo $N(k,t) \to N(k,t) + 1$

O número esperado de links que vai conectar com um nó de grau K é

$$\underbrace{\frac{k}{2mt}}_{\Pi(k)} \times \underbrace{N P_k(t)}_{\substack{\text{no de nós} \\ \text{com grau } k \\ (N = t)}} \times \underbrace{m}_{\substack{\text{no de links} \\ \text{do} \\ \text{novo nó}}} = \frac{k}{2} P_k(t)$$

Temos então a seguinte "equação mestra":

$$(N+1)\, P_k(t+1) = N P_k(t) + \frac{(k-1)}{2} P_{k-1}(t) - \frac{k}{2} P_k(t)$$

$(N+1)\,P_k(t+1)$: Nº de nós com grau k em t+1

$N P_k(t)$: Nº de nó com grau k em t

$\frac{(k-1)}{2} P_{k-1}(t)$: Nº de nós com k−1 em t que ganhou um link

$\frac{k}{2} P_k(t)$: Nº de nós com k em t que ganhou um link e ficou com (k+1)

Como cada nó chega com grau m, não existem nós com grau (m−1) que podem ganhar um link e virar grau m. Então, a equação acima só vale p/ $k > m$. Para $k = m$ o segundo termo do lado direito é igual a 1, que representa o novo nó.

$$(N+1)\, P_m(t+1) = N P_m(t) + 1 - \frac{m}{2} P_m(t)$$

Quando $t = N \to \infty$ a distribuição atinge um equilíbrio. Nesse limite

$$(N+1) P_k(t+1) - N P_k(t) = (N+1) P_k(\infty) - N P_k(\infty) = P_k(\infty)$$

$$(N+1) P_m(t+1) - N P_m(t) = (N+1) P_m(\infty) - N P_m(\infty) = P_m(\infty)$$

Vamos denotar $P_k(\infty)$ apenas por P_k. Então

$$\begin{cases} P_k = \frac{k-1}{2} P_{k-1} - \frac{k}{2} P_k \\ P_m = -\frac{m}{2} P_m + 1 \end{cases}$$

$$P_m \left(1 + \frac{m}{2}\right) = 1 \quad \Rightarrow \quad \boxed{P_m = \frac{2}{m+2}}$$

$$P_k \left(1 + \frac{k}{2}\right) = \frac{k-1}{2} P_{k-1} \quad \Rightarrow \quad P_k = \frac{k-1}{k+2} P_{k-1} \quad , \ k > m$$

ou $$\boxed{P_{k+1} = \frac{k}{k+3} P_k} \qquad k \geq m-1$$

Solução

$$P_m = \frac{2}{m+2}$$

$$P_{m+1} = \frac{m}{m+3} P_m = \frac{2m}{(m+2)(m+3)}$$

$$P_{m+2} = \frac{m+1}{m+4} P_{m+1} = \frac{2m(m+1)}{(m+2)(m+3)(m+4)}$$

$$\vdots$$

$$P_{m+3} = \frac{m+2}{m+5} P_{m+2} = \frac{2m(m+1)}{(m+3)(m+4)(m+5)}$$

$$\vdots$$

$$\boxed{P_k = \frac{2m(m+1)}{k(k+1)(k+2)}}$$

- p/ $k \gg 1$ $\quad P_k \sim k^{-\gamma}$ com $\gamma = 3$
- γ é independente de m
- P_k é independente de N (ou t), para t grande
- coeficiente correto é $m(m+1)$, não m^2.

RE-OBTENDO P(K) da equação MESTRA

Da página (50) vimos que

$$P_k = \frac{k-1}{2} P_{k-1} - \frac{k}{2} P_k$$

$$2 P_k = (k-1) P_{k-1} - k P_k = - P_{k-1} - k \left(P_k - P_{k-1} \right)$$

tomando o limite de k contínuo podemos escrever

$$2P(k) = -P(k-1) - k \left(\frac{P(k) - P(k-1)}{k - (k-1)} \right)$$

ou

$$2P(k) = -P(k) - k \frac{dP}{dk} = - \frac{d}{dk} \left(k P(k) \right)$$

e

$$\boxed{P(k) = -\frac{1}{2} \frac{d}{dk} \left(k P(k) \right)}$$

A solução, voltando um passo, é

$$3 P(k) = -k \frac{dP}{dk} \Rightarrow -3 \frac{dk}{k} = \frac{dP}{dP} \quad -3 \ln k = \ln P$$

$$P(k) \simeq k^{-3} \quad //.$$

O PAPEL DA CONEXÃO Preferencial

O que ocorre se um rede cresce sem conexão preferencial.?

Se a ligação dos novos nós é aleatória, então

$$\Pi_i = \frac{1}{N} = \frac{1}{m_0 + t - 1}$$

e

$$\frac{dk_i}{dt} = m \Pi_i = \frac{m}{m_0 + t - 1}$$

$$k_i(t) = m \left[\ln \left(\frac{m_0 + t - 1}{m_0 + t_i - 1} \right) + 1 \right] \rightarrow \text{crescimento lento}$$

Se $t_i < (m_0 - 1 + t)\, e^{-\frac{k}{m}+1} + 1 - m_0$ então $k_i > k$

de nós com $k > k_i$ é $t_i = (m_0 - 1 + t)\, e^{-\frac{k}{m}+1} + 1 - m_0$

$$P(k) = 1 - t_i/t = 1 - \frac{m_0 - 1 + t}{t} e^{-k/m+1} - \frac{(1-m_0)}{t}$$

$$p(k) = \frac{dP}{dk} = \frac{m_0 - 1 + t}{m\,t} e^{-k/m+1} \simeq \frac{e}{m} e^{-k/m}$$

que é uma distribuição exponencial.

A conexão preferencial tem papel fundamental no aparecimento da lei de potência.

O papel do crescimento

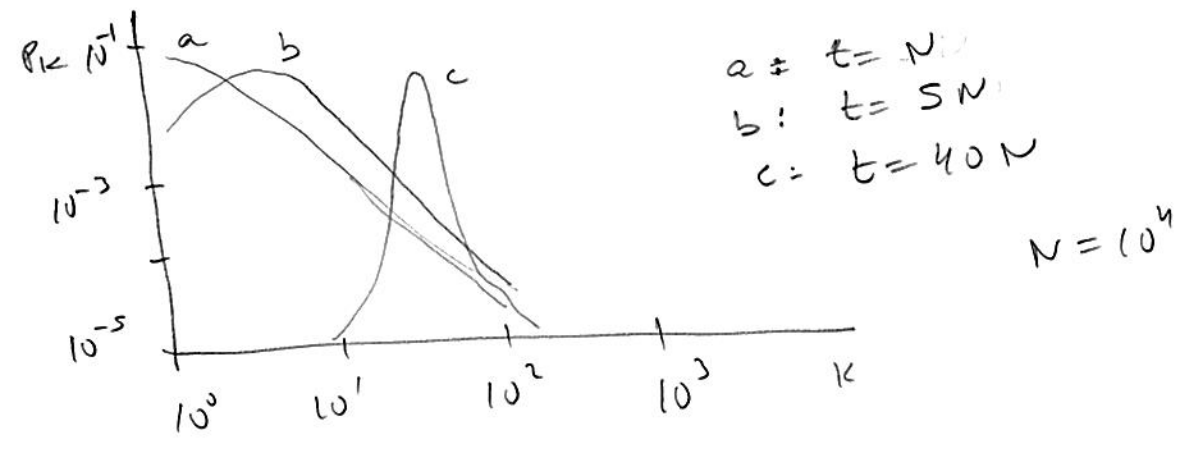
Vamos agora considerar uma rede com N nós e conectá-los sorteando um nó de cada vez e fazendo a ligação com a conexão preferencial. Com isso testamos o que acontece se a rede não cresce. Nós com $K=0$ são supostos ter $K=1$ p/ o cálculo de Π_i.

N é fixo
nº de links aumenta linearmente, 2 por passo de tempo

Então $k_i(t) = \frac{2}{N} t$ onde

$\frac{1}{N}$ = prob. do nó i ser sorteado
$2t$ = nº de links no tempo t

Se t é pequeno e o número de links $L << N$ a rede se comporta como BA. No entanto $P(k)$ não fica estacionária e tende à um grafo completo:



VI- Fitness: o modelo de Bianconi-Barabasi

Nem sempre o número de conexões de um nó é o principal atrativo para novos nós se conectarem. Muitas vez os nós exibem alguma qualidade que os torna mais pré-dispostos a atraírem conexões. Exemplos

- em um voo internacional as opções de escala são feitas por distância ao ponto de destino e pela qualidade do aeroporto (facilidade de mudar de terminal, etc).

- predadores com dieta generalista vão se conectar facilmente a uma presa invasora.

- o facebook passou o google, mesmo tendo aparecido depois.

Modelo: - η_i = qualidade do nó i

- Para cada novo site η é escolhido de uma distribuição $\rho(\eta)$.

- probabilidade de conexão é

$$\Pi_i = \frac{\eta_i K_i}{\sum \eta_j K_j}$$

Exemplos de $\rho(\eta)$

(a) $\rho(\eta) = \delta(\eta - 1)$ → todos os sítios tem $\eta = 1$

(b) $\rho(\eta) = 1$ p/ $0 < \eta < 1$

→ distribuição uniforme

(c) $$\rho(\eta) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(\eta-\eta_0)^2/2\sigma^2}$$

→ GAUSSIANA centrada em η_0 com VARIÂNCIA σ

Note que todos satisfazem $\int \rho(\eta) d\eta = 1$

A evolução temporal do grau do nó i, k_i, segue a fórmula

$$\frac{\partial k_i}{\partial t} = m \Pi_i = \frac{m k_i \eta_i}{\sum k_j \eta_j}$$

Vamos mostrar que essa equação pode ser resolvida se considerarmos médias sobre diversas realizações da rede, ou seja, médias sobre a distribuição $\rho(\eta)$.

Os resultados são os seguintes:

$$K_i(t, t_{0i}, \eta_i) = m \left(\frac{t}{t_{0i}} \right)^{\beta(\eta_i)}$$

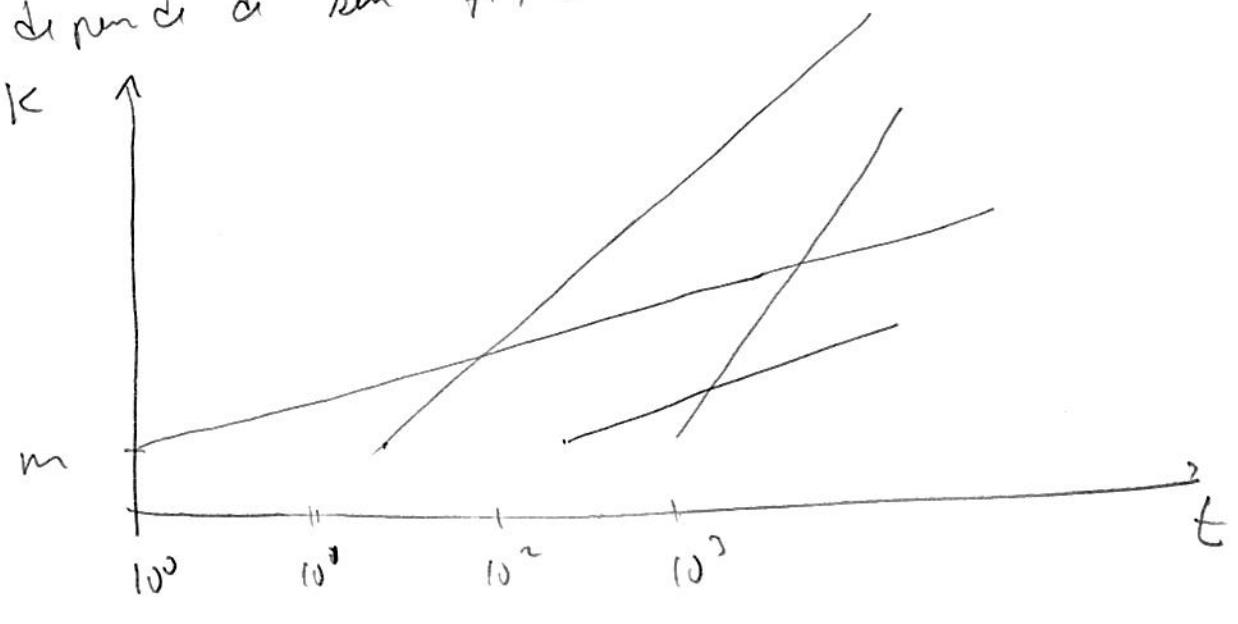
η_i = fitness sorteado p/ o nó i

t_{0i} = instante que i entra na rede

$$\beta(\eta_i) = \frac{\eta_i}{C}$$

$$C = \int \frac{\eta\, \rho(\eta)}{1-\beta(\eta)} d\eta = \int \frac{\eta\, \rho(\eta)}{1-\eta/C} d\eta$$

Assim, dada a distribuição $\rho(\eta)$ podemos calcular o valor de C e com isso teremos o expoente de cada nó $\beta_i = \eta_i / C$. Cada nó cresce com expoente que depende de seu fitness:



A distribuição de grau também pode ser calculada e é dada por

$$P_k \approx \frac{C}{m} \int d\eta \, \frac{\rho(\eta)}{\eta} \left(\frac{m}{k}\right)^{\frac{C}{\eta}+1}$$

e é um composição de Leis-de-potência $\left(\frac{m}{k}\right)^{\beta+1}$ cujos pesos são dados por $\frac{\rho(\eta)}{\eta}$.

Exemplos :

(1) Todos os nós tem o mesmo fitness $\eta = 1$

$$\rho(\eta) = \delta(\eta - 1)$$

onde a função "delta de Dirac" tem as propriedades

$$\int \delta(\eta - \eta_0) \, d\eta = 1$$

$$\int f(\eta) \, \delta(\eta - \eta_0) \, d\eta = f(\eta_0)$$

Usando essa segunda propriedade podemos calcular C :

$$C = \int \left(\frac{\eta}{1 - \eta/C}\right) \delta(\eta - 1) \, d\eta = \frac{1}{1 - 1/C} \Rightarrow \boxed{C = 2}$$

$$\beta(1) = \frac{1}{c} = \frac{1}{2}$$

$$P_k = \frac{2}{m}\int dn \left[\left(\frac{m}{k}\right)^{\frac{2}{n}+1} \frac{1}{n}\right] \delta(n-1)$$

$$= \frac{2}{m}\left(\frac{m}{k}\right)^{2+1} = .2\, m^2 k^{-3}$$

que corresponde ao resultado Barabasi- Albert

(2) Distribuição uniforme no intervalo $[0,1)$

$$g(n) = 1 \quad , \quad \int_0^1 g(n)\, dn = 1$$

$$C = \int_0^1 \frac{n}{1-n/c}\, dn = c\int_0^1 \left[\frac{1}{1-n/c} - 1\right] dn = c\left[-c\ln(1-1/c) - 1\right]$$

$$1 = -c\ln(1-1/c) - 1 \quad \Rightarrow \quad -2/c = \ln(1-1/c)$$

ou $\boxed{1 - 1/c = e^{-2/c}}$

A solução pode ser obtida numericamente e resulta

$$c^* = 1.255$$

$$\beta(n_i) = \frac{n_i}{c^*}$$

$$P_K = \frac{c^*}{m}\int_0^1 \frac{d\eta}{\eta}\left(\frac{m}{K}\right)^{\frac{c^*}{\eta}+1} \simeq \frac{K^{-(c^*+1)}}{\ln K}$$

$$\simeq \frac{K^{-2.255}}{\ln K}$$ (veja pag. R62)

DEMONSTRAÇÃO DOS RESULTADOS

Para resolver a equação

$$\frac{\partial k_i}{\partial t} = \frac{m\, k_i\, \eta_i}{\sum k_j \eta_j}$$

calculamos o valor médio da soma no denominador em relação à várias réplicas da mesma rede..

Em cada réplica associamos um valor de η_i tirado de $\rho(\eta)$ a cada nó. Fixando um valor de η podemos nos perguntar qual nó será associado a esse fitness. Na primeira réplica pode ser o 10º nó, na segunda o 25º nó, etc..

Em cada réplica, o nó que fica com η aparece em um t_0 diferente. Podemos então aproximar

$$\left\langle \sum_j \eta_j k_j \right\rangle = \int d\eta\, \rho(\eta)\, \eta \int_1^t dt_0\, K(t, t_0, \eta)$$

Substituindo

$$K = m\left(\frac{t}{t_0}\right)^{\beta(\eta)} \quad \text{obtemos}$$

$$\left\langle \sum_j \eta_j k_j \right\rangle = \int d\eta\, \rho(\eta)\, \eta\, m\, t^{\beta} \underbrace{\int_1^t t_0^{-\beta}\, dt_0}_{\frac{t^{-\beta+1}-1}{1-\beta}}$$

$$= \int d\eta\, \rho(\eta)\, m\, \eta \left(\frac{t - t^{\beta}}{1-\beta}\right)$$

Usamos agora o fato que $0 < \beta < 1$. De fato, $\beta > 0$ pois o grau de um nó só pode crescer com o tempo. Além disso a taxa máxima de crescimento é mt, pois apenas um nó entra na rede por passo de tempo. Para $t \gg 1$ (redes grandes)

$$t^{\beta} \to 0 \quad \text{e}$$

$$\left\langle \sum_j \eta_j k_j \right\rangle = m\,t \int \frac{d\eta\, \eta\, \rho(\eta)}{1-\beta(\eta)} \equiv m\,t\,C$$

Então

$$\frac{\partial k_i}{\partial t} = \frac{m \eta_i k_i}{\sum \eta_j k_j} = \frac{m \eta_i k_i}{m t C} = \frac{\eta_i k_i}{t C}$$

cuja solução de fato é

$$k_i = m\left(\frac{t}{t_0}\right)^{\eta_i / C} \Rightarrow \beta(\eta) = \frac{\eta}{C}$$

Finalmente a integral na pág. 1260 pode ser calculada da seguinte forma:

$$P_k = \frac{C^*}{m}\int \frac{d\eta}{\eta}\left(\frac{m}{k}\right)^{\frac{C^*}{\eta}+1} \quad ; \quad \frac{C^*}{\eta} = x \rightarrow -\frac{C^*}{\eta^2}d\eta = dx \rightarrow d\eta = -\frac{C^*}{x^2}dx$$

$$a \equiv k/m$$

$$= \frac{1}{m}\int_{C^*}^{\infty}\left(\frac{C^*}{x^2}\right)dx \times a^{-x-1} = \frac{C^*}{k}\int_{C^*}^{\infty}\frac{dx}{x}\, e^{-x \ln a}$$

$$\approx \frac{C^*}{k}\int_{C^*}^{\infty}\frac{dx}{C^*}\, e^{-x \ln a} = \frac{1}{k \ln a}\, e^{-C^* \ln a} = \frac{1}{k \ln a}\, a^{-C^*}$$

$$= \frac{k^{-(C^*+1)}}{(\ln k - \ln m)}\, m^{C^*} \approx \frac{m^{1.255}}{\ln k}\, k^{-2.255}$$