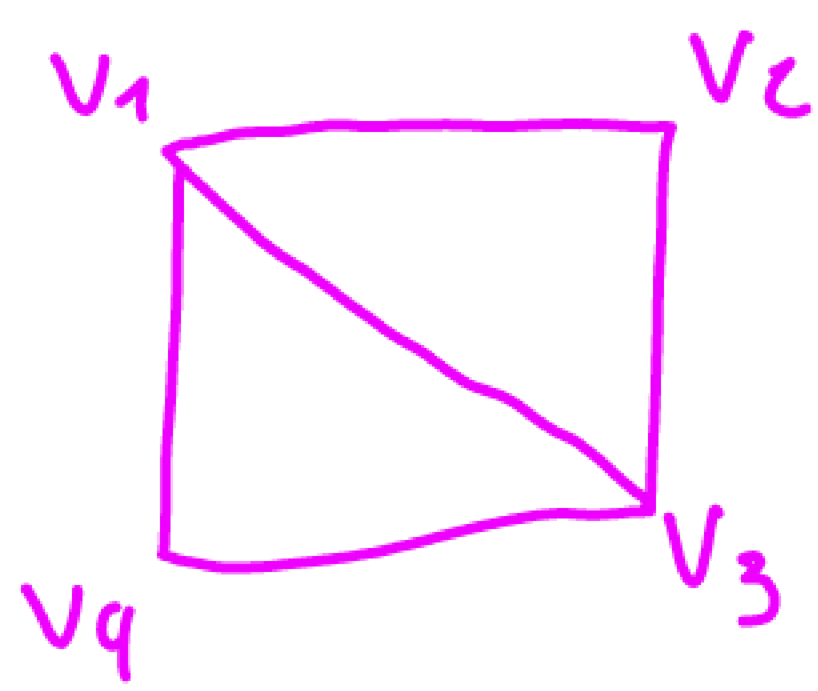
By Małgorzata Sulkowska, the last one on PWR.

Complex Networks

Def: A simple graph G is a pair (V, E)
where V - set of vertices
E - set of edges (2-element subsets of V)

example:

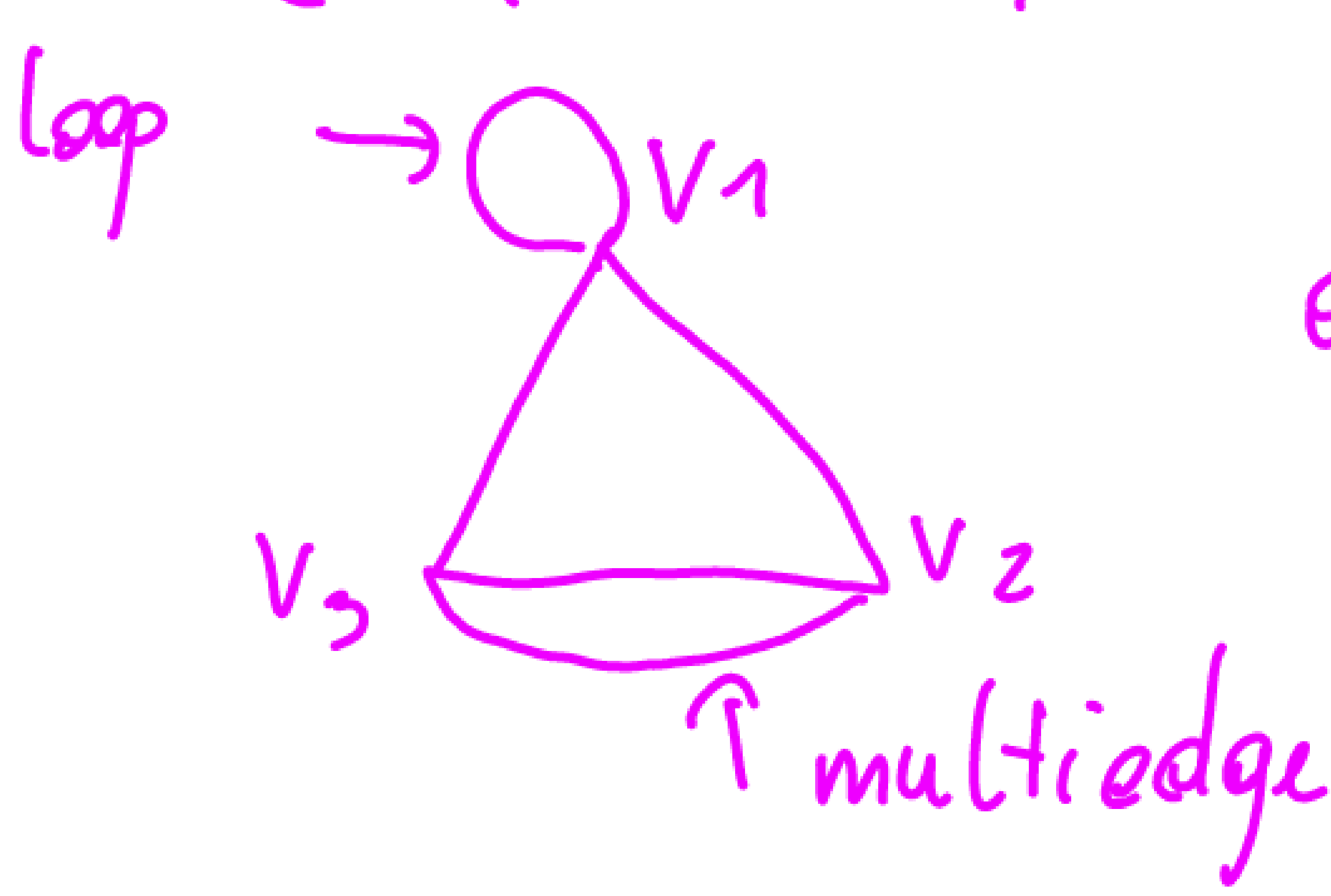


$\deg(v_3) = 3$

Def: A (multi)graph G is a pair (V, E), where
V - set of vertices
E - set of edges (1-, 2- element subsets of V)

example:

$G = \{\{v_1, v_2, v_3\}, \{\{v_1, v_2\}, \{v_2, v_3\}, \{v_2, v_3\}, \{v_3, v_1\}, \{v_1\}\}$



$\deg(v_1) = 3$ // loop $\equiv +1$
$\deg(v_2) = 3$

generally $\deg(v) \equiv$ number of edges integrated with v.

Def: Complex network – a large graph that reflects well real-life system (the topology of this graph is complicated enough)

example: of r-l systems:
- social network
- protein (biological network)
- transportation networks
- technological network

- Barabasi, Albert (1999) Huge step
  preferential attachment model
  (rich-get-richer, poor-get-poorer)
  this describes well the "growth" of graph's life.

- Chung, Lu (2006)
  extended preferential attachment model.

parameters:
- probability $p \in [0,1]$
- artificial graph G

we construct a sequence of graphs $\{G_t\}_{t \geq 0}$:
- $G_0 \equiv$ (1 vertex + 1 loop)
- G_{t+1} is built upon G_t.
  - with probability p take "vertex-step":

$G_{t+1} \equiv$ 

$$P[u \text{ is chosen}] = \frac{\deg(u)}{\sum_{w \in G_t} \deg(w)}$$

  - otherwise take "edge-step":

$G_{t+1} \equiv$ 

u, v – chosen with probability proportional to their degree (possible with repetition)

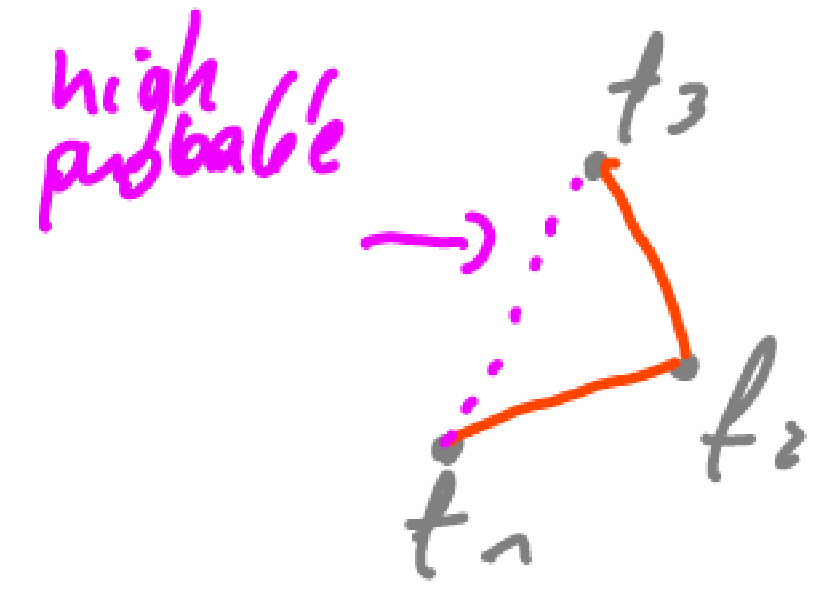
weights are updated after the step

Early 2000s, real-life systems feature:

1. Small diameter (the length of the shortest path between the most distanced vertices)
(small world phenomen)
(six-degree of seperation)

2. High clustering coefficient $\sim \dfrac{\#\Delta\text{'s}}{\#\Lambda\text{'s}}$
triangle
cherry

(friend of my friend is likely to become my friend)

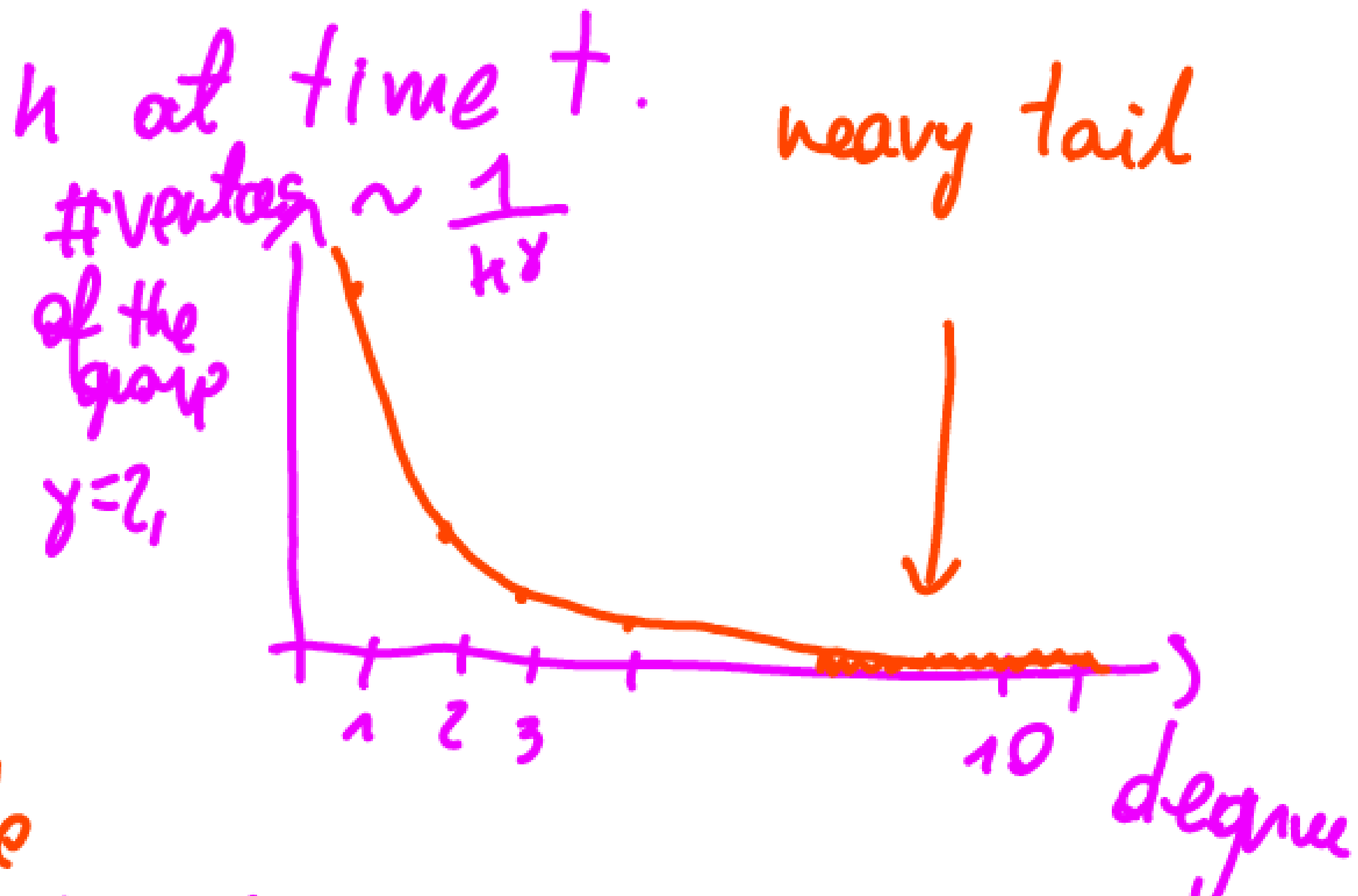


3. Visible community structure.

modularity - parameter to measure this structure.

4. Power-law degree distribution.
$M_{k,t}$ - # of vertices of degree k at time t.

$$\frac{M_{k,t}}{|V_t|} \sim C \cdot k^{-\gamma}, \quad \gamma > 1$$



Power-law has dist like very rich people

$1: \sim 1$

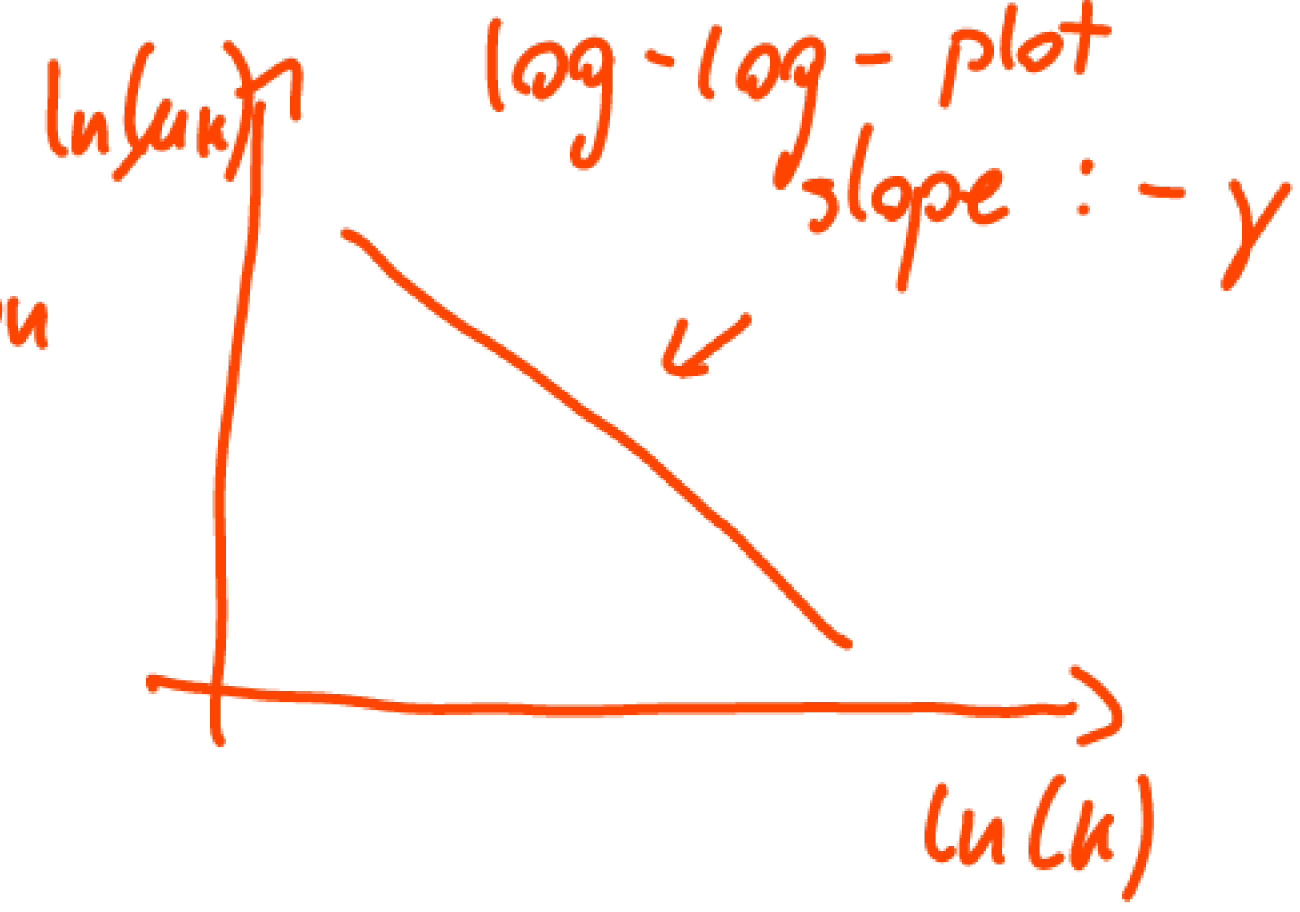
$2: \sim \frac{1}{2^2}$

$3: \sim \frac{1}{3^2}$

$4: \sim \frac{1}{4^2}$

let $\mu_k = \frac{M_{k,t}}{|V_t|}$

then $\ln(\mu_k) \sim \gamma \cdot \ln(k) + \ln(C)$ then



---

going back to theoritical model: V - 1,4 , X - 2,3

in G_t:

$|E_t| = t+1$

$|V_t| = 1 + X_t \qquad X_t \sim Bin(t, p)$

$E[V_t] = 1 + p \cdot t$, $|V_t|$ - "highly concentrated", chernoff inequality

$M_{k,t}$ - # of vertices of degree k at time t.

aim:

$\lim\limits_{t\to\infty} E\left[\frac{M_{k,t}}{|V_t|}\right] = ?$ $\xrightarrow{\text{given } V_t \text{ is}}$ $\lim\limits_{t\to\infty} \frac{E[M_{k,t}]}{p t} = ? \overset{?}{\sim} C \cdot k^{-\gamma}$

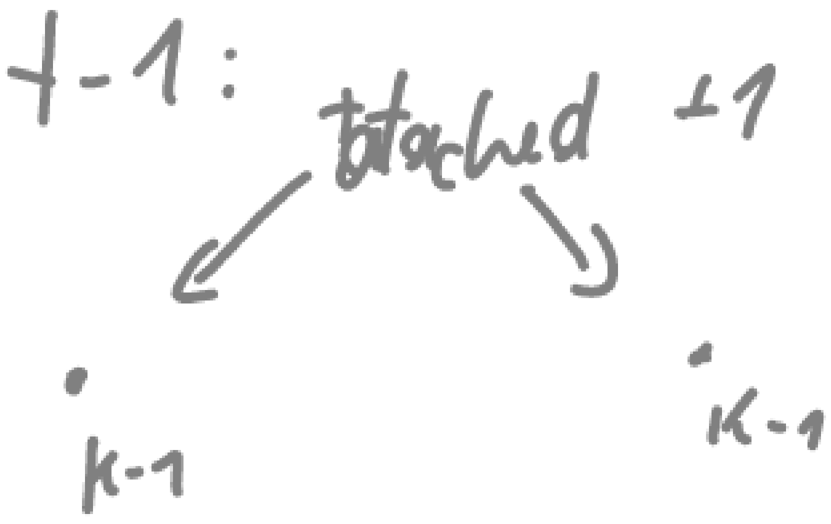
Want to derive the recurrence formula for $E[M_{k,t}]$.

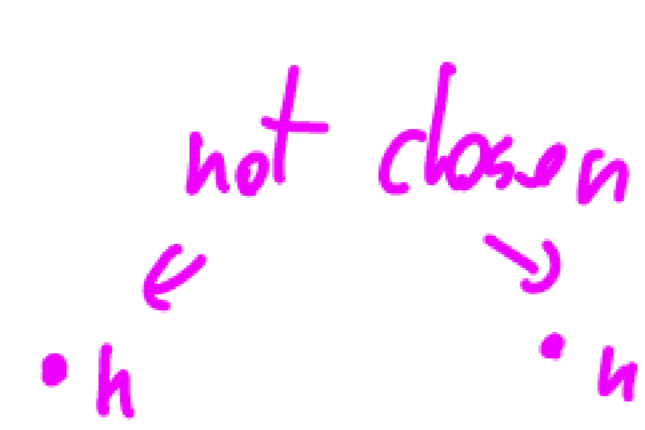
$\mathcal{F}_t$ - σ-algebra associated with probability space at time t.

$E[M_{k,t} \mid \mathcal{F}_{t-1}] =$ we know all info before $M_{k,t}$ for example $M_{k,t-1}$, $M_{k-1,t-1}$

$$p\left[M_{k-1,t-1}\left(\frac{k-1}{2t}\right) + M_{k,t-1} \cdot \left(1 - \frac{k}{2t}\right)\right] + (1-p)\left[M_{k-1,t-1}\left(\frac{2(k-1)}{2t} - \left(\frac{k-1}{2}\right)^2\right) + M_{k,t-1} \cdot \left(1 - \frac{2k}{2t} + \left(\frac{k}{2t}\right)^2\right)\right] = \dots \to$$

in t-1:





$$\ldots = M_{k,t-1} \cdot \left[1 - \frac{(2-p)k}{2t} + (1-p)\left(\frac{k}{2t}\right)^2 \right] +$$

$$+ M_{k-1,t-1} \cdot \left[\frac{(2-p)(k-1)}{2t} - (1-p) \cdot \left(\frac{k-1}{2t}\right)^2 \right]$$

btw: $E[E[X|Y]] = E[X]$
so when we put in on both sides

$$E[M_{k,t}] = E[M_{k,t-1}] \cdot [\ldots] + E[M_{k-1,t-1}] \cdot [\ldots]$$

now, the case $\forall t > 0, k = 1$

- $\underbrace{E[M_{1,t}]}_{a_{t+1}} = \underbrace{E[M_{1,t-1}]}_{a_t}\left(1 - \frac{2-p}{2t} + (1-p) \cdot \left(\frac{1}{2t}\right)^2\right) + p \cdot 1$ ($\frac{b_t}{t}$, c_t)
- $E[M_{1,0}] = 1$ recurrence equation

Lemma 1   Let $a_{t+1} = \left(1 - \frac{b_t}{t}\right) \cdot a_t + c_t$, $\lim_{t\to\infty} b_t = b$, $\lim_{t\to\infty} c_t = c$

then $\lim_{t\to\infty} \frac{a_t}{t} = \frac{c}{1+b}$

I $\lim_{t\to\infty} b_t = \lim_{t\to\infty} \left(\frac{2-p}{2} + \frac{(1-p)}{4t}\right) = \frac{2-p}{2}$, II $\lim_{t\to\infty} c_t = p$

$$\lim_{t\to\infty} \frac{E[M_{1,t}]}{t} = \frac{p}{1+\frac{2-p}{2}} = \frac{2p}{4-p}$$

How about $k > 1$?
induction on k. Assume $M_{k-1} := \lim_{t\to\infty} \frac{E[M_{k,t}]}{t}$ exists
and apply Lemma 1 to $(**)$. we'll get

$$\begin{cases} M_k = M_{k-1} \cdot \frac{(k-1)}{k + \frac{2}{2-p}} \\ M_1 = \frac{2p}{4-p} \end{cases} \Rightarrow M_k = \frac{2p}{4-p} \cdot \frac{1 \cdot 2 \cdot \ldots \cdot (k-1)}{\left(2 + \frac{2}{2-p}\right)\left(3 + \frac{2}{2-p}\right) \ldots \left(k + \frac{2}{2-p}\right)} =$$

$$= \frac{2p}{4-p} \cdot \frac{\Gamma(k) \cdot \Gamma\left(2+\frac{2}{2-p}\right)}{\Gamma\left(k+\frac{2}{2-p}+1\right)}$$

<u>Lemma 2</u>

$\alpha \in \mathbb{R}$ $\quad \lim\limits_{n\to\infty} \frac{\Gamma(n) \cdot n^{\alpha}}{\Gamma(n+\alpha)} = 1$, $\frac{\Gamma(n)}{\Gamma(n+\alpha)} \sim n^{-\alpha}$

$$\sim \frac{2p}{4-p} \cdot \Gamma\left(2+\frac{2}{2-p}\right) \cdot k^{-\left(1+\frac{2}{2-p}\right)}$$

and that is it.