

Spring - 1959

TELEPHONE: CENTRAL 2266

Cable address: OSTERPORTHOTEL Copenhagen



HOTEL ØSTERPORT

(1)

COPENHAGEN, DENMARK

Dear Bob

While drinking beer yesterday several ideas came about our maximization problem and the question of root selection. You have probably solved the problem by now, but here are a few points anyhow. First, as you know, the solution is a simple saddle point of $H = \sum_i i(1 - e^{-\lambda_i x_i}) - \lambda \sum x_i + \mu \sum j_i$, gotten by setting both partials to zero. The problem is how to select between the 3 roots which sometimes occur. I have thought a little about why the "magic" multipliers (Lagrange) work to express constraints, and will try to briefly give you the picture: If we have a function $f(x_i)$ to be minimized subject to constraints $f_j(x_i) = c_j$, then we can picture the constraints as limiting our region of interest to a "surface" (subspace) within the x_i space. There is in fact a whole family of surfaces which can be labeled by the values of the $f_j(x_i)$ which are constant over one surface.

If we now choose some fixed values λ_1, λ_2 in



HOTEL STEFERTS

(5)

COPENHAGEN, DENMARK

and maximize (by any method, not necessarily just setting derivatives to zero) the expression $f(x) - \lambda_1 g_1(x) - \lambda_2 g_2(x) \dots$ (without further constraints except for possibly positivity), then this maximum will in fact be a maximum for the constrained problem for some set of constraint values λ_i . The reason for this is: wherever the maximum lies, it is on some surface, i.e. the surface of constant $g_i(x)$ which passes thru the point calculated by maximizing the expression with the multipliers. The crucial point is that it is also a maximum relative to all other points on the constraint surface, since the constraint terms $\lambda_i g_i(x)$ are constant over the whole surface — so that a maximum of $f(x) - \lambda_i g_i(x) \dots$ implies an absolute maximum of $f(x)$ alone within that surface! Therefore simple maximization of $f(x) - \lambda_i g_i(x) \dots$ for any choice of $\{\lambda_i\}$ leads to a solution which is a maximum for some constraint set. The actual dependence of the λ_i 's on the constraint values c_j is, of course, an induced dependence, and can be gotten only by solving the maximization for all λ_i 's and then calculating c_j 's.



HOTEL ØSTERPORT

(3)

COPENHAGEN, DENMARK

So much for the theory of Lagrange multipliers. The main point for us is that the magic is not in the derivatives but simply in the maximization. Player 2 is merely interested in maximizing

$H = \sum_i h_i (1 - e^{-A_i n_i}) - \lambda \sum_i n_i + \mu \sum_i j_i$, which (an important point) amounts to independently minimizing each

$h_i = \frac{1}{n_i} (1 - e^{-A_i n_i}) - \lambda n_i + \mu j_i$ -- and the problem really separates to independent consideration for each i !

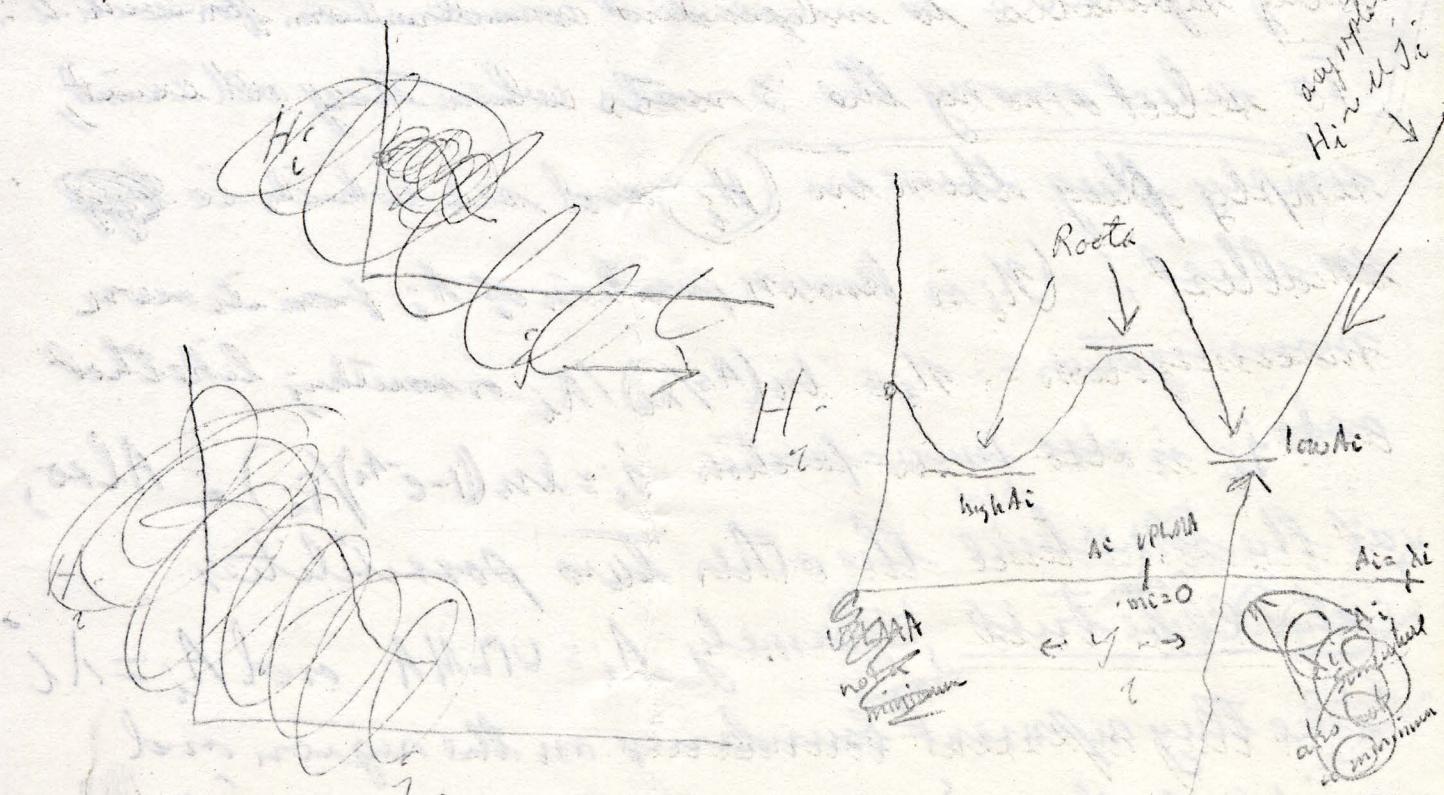
To select among the 3 walls when they all exist, simply plug them in (h_i) and see which is ~~biggest~~ smallest. (n_i is known function of A_i from its own maximization -- no $\ln(A_i/n_i)/A_i$ or something like that and j_i is also known function $j_i = \ln((1 - e^{-A_i})/p_i)$). Also, ~~at the same time~~ the other two possibilities ~~would be tried~~, namely $A_i = \text{UPMA}$ and $A_i = \bar{A}_i$ since they represent boundaries on the region and might be the maxima. So there is a fairly simple selection criterion.



НЕТОН

STERPORT

(4) Now, one can probably do a little better than just trying all 5 (or 3) possibilities each time. A little analytic work with H_i and some fiddling of inequalities might show that always the lower (or upper) root was correct. (Actually I doubt this.) Another point, however, is that the middle root is not a solution, but must correspond to a local maximum for purely topological reasons.



Since for $i \rightarrow \infty$ H_i is ∞ , the first turning point for decreasing i can only be a minimum, hence the next only a maximum, and the last again a minimum. Hence middle root not necessary.



HOTEL ØSTERPORT

(5)

COPENHAGEN, DENMARK

I hope that these considerations are not too late to be of some help. Anyhow a better understanding of real meaning of Lagrange Multipliers should be helpful in future problems of this type. Please drop me a note at above address and let me know what's going on. I'll be here till about 24 April.

Good Luck
Hugh

P.S. The choice of signs $- \lambda \Sigma j_i$ and $+\mu \Sigma j_i$ in the expression is purely so that λ and μ will turn out to be positive numbers for actual solutions. Of course they could be entered with opposite signs in the expression, but then would have turned out to be negative for meaningful solutions, and the above arguments would all still hold (argument about exclusion of middle well) of course.