4.2

Maximal Solutions of the Generalized Subadditive Inequality

J. M. Hammersley and G. R. Grimmett

4.2.1 INTRODUCTION

In an enquiry on maximum likelihoods of genetic histories, Harding [1] met a function defined recursively for $n = 1, 2, \ldots$ by means of

$$(1) \qquad M(1) = 1, \quad M(n) = \max_{1 \leqslant r \leqslant \frac{1}{2}n} \left[\frac{2M(r)M(n-r)}{(n-1)} \right] \quad (n \geqslant 2).$$

Writing $\rho(n)$ for the value of r which maximized the right-hand side of equation (1) he conjectured that either $\rho(n)$ or $n-\rho(n)$ was always a power of 2 lying between $\frac{1}{3}n$ and $\frac{2}{3}n$; and he verified this conjecture on a computer for $n \leqslant 100$. More specifically, his conjecture amounts to

$$(2) \qquad \rho(n) = \begin{cases} 2^{k-1} & \text{if } 1 \leqslant n/2^k \leqslant \frac{3}{2}, \\ n-2^k & \text{if } \frac{3}{2} \leqslant n/2^k \leqslant 2. \end{cases}$$

We shall prove this conjecture in the wider setting (with other applications, perhaps) of *generalized subadditive functions* [2], namely sequences

$$\varphi = \{\varphi(n)\}_{n=1,2,\ldots}$$

satisfying the inequality

$$(3) \qquad \varphi(n) \leqslant \varphi(r) + \varphi(n-r) + g(n) \quad (1 \leqslant r < n \geqslant 2),$$

where $g = \{g(n)\}_{n=1,2,\ldots}$ is a given sequence. If $\varphi(n)$ is a solution of (3), so is $\varphi(n) - \lambda n$ for any constant λ; and hence, taking $\lambda = \varphi(1)$, we may confine ourselves to the *canonical solution* of (3), namely the solution with $\varphi(1) = 0$. We may also assume $g(1) = 0$, because (3) only involves $g(n)$ for $n \geqslant 2$. We shall write S for the space of all real sequences $\{s(n)\}_{n=1,2,\ldots}$ with $s(1) = 0$. Thus we are looking for solutions $\varphi \in S$ of (3) for given $g \in S$.

We can regard S as a complete topological space over the reals, with vector addition $\alpha s + \beta t = \{\alpha s(n) + \beta t(n)\}$ as usual, and some convenient metric, say

$$(4) \qquad d(s,t) = \lim_{n\to\infty} \frac{\Sigma_n}{1+\Sigma_n}, \quad \Sigma_n = \sum_{i=1}^{n} |s(i) - t(i)|.$$

Indeed S only fails to be a Banach space by lacking $d(\alpha s, 0) = |\alpha| d(s, 0)$.

The inequality (3) is equivalent to

$$(5) \qquad \varphi(n) \leqslant \min_{1 \leqslant r \leqslant \frac{1}{2}n} [\varphi(r) + \varphi(n-r)] + g(n) \quad (n \geqslant 2).$$

Hence, if we define $f \in S$ recursively by means of

$$(6) \qquad f(1) = 0, \quad f(n) = \min_{1 \leqslant r \leqslant \frac{1}{2}n} [f(r) + f(n-r)] + g(n) \quad (n \geqslant 2),$$

then f is a solution of inequality (3). Moreover it is the maximal solution of (3) in the sense that $f(n) \geqslant \varphi(n)$ for all n and any solution φ of inequality (3). This follows at once by induction on n. Harding's relation (1) is equivalent to equation (6) in the particular case

$$(7) \qquad f(n) = -\log M(n), \quad g(n) = \log \tfrac{1}{2}(n-1).$$

We propose to study equation (6) in general, and we define $\sigma(n) = \sigma(n,f)$ to be the set of integers r in $1 \leqslant r \leqslant \frac{1}{2}n$ which minimize $f(r) + f(n-r)$. Let $\rho = \{\rho(n)\}_{n=2,3,\ldots}$ be any given sequence of integers such that $1 \leqslant \rho(n) \leqslant \frac{1}{2}n$, and define S_ρ to be the set of all $f \in S$ such that $\sigma(n,f) = \rho(n)$ for $n = 2, 3, \ldots$. (*Note:* we regard a sequence as a one-valued function of n; so this definition is framed to exclude f from every S_ρ if the set $\sigma(n,f)$ contains more than one member for some n.) Any S_ρ is either empty or else a convex positive cone of S, and S is the union (over all ρ) of the closures of the S_ρ.

[To sketch a proof of the closure assertion, suppose $f \in S$ and prescribe $\varepsilon > 0$. Let $h = \{h(i)\}$ be a sequence of independent random variables $h(i)$, with $h(i)$ uniformly distributed on the interval $[f(n) - 2^{-n}\varepsilon, f(n) + 2^{-n}\varepsilon]$. Then, with probability 1, $\sigma(n,h)$ is single-valued for all $n \leqslant n_0$ for any prescribed n_0, and hence for all n. Since any such sequence h satisfies $d(f,h) < \varepsilon$, there exists certainly (not merely with probability 1) some sequence $h^* \in \bigcup_\rho S_\rho$ such that $d(f, h^*) < \varepsilon$.]

From equation (6) we see that f is a continuous transform of g, which we may write

$$(8) \qquad f = Tg.$$

Indeed, recursive solution of equation (6) will yield $f(n)$ as a linear combination

(9) $$f(n) = \sum_{m \leqslant n} t_{nm} g(m),$$

where the t_{nm} are integers.

The inverse transform

(10) $$g = T^{-1} f$$

is also continuous and, when $f \in S_\rho$, can be written as

(11) $$g(n) = f(n) - f[\rho(n)] - f[n - \rho(n)].$$

However, the transformation T is not linear on S. It is merely positive-homogeneous

(12) $$T(\alpha g) = \alpha T g \quad (\alpha \geqslant 0)$$

and superadditive

(13) $$T(g+h) \geqslant Tg + Th,$$

because $Tg + Th$ is a solution of inequality (3) with $g(n) + h(n)$ in place of $g(n)$. In matrix form T and T^{-1} are both lower triangular with unit diagonal.

While the foregoing exhibits the general structure (in fashionable but equally superficial jargon), the heart of our problem lies in specifying the appropriate sequence ρ. Specifically, given g, to which region S_ρ, if any, does Tg belong? Once this question is answered, we can compute f from (9). We have only answered the question in the three simplest cases:

Case (i): If g is strictly decreasing, then $\rho(n) = 1$.

Case (ii): If g is increasing and strictly convex, then $\rho(n) = [\frac{1}{2}n]$, the integer part of $\frac{1}{2}n$.

Case (iii): If g is increasing and strictly concave, then $\rho(n)$ is given by equation (2).

Proofs appear in the next sections, together with the corresponding explicit solutions (20), (27) and (41) for f.

There are sequences g such that Tg belongs to no S_ρ. (For example, if $g = (0, 0, \ldots)$ then $f = Tg = (0, 0, \ldots)$ and $\sigma(n)$ clearly consists of all integers in $[1, \frac{1}{2}n]$. Thus $\sigma(n)$ is not single-valued and Tg is a limit point of every S_ρ but a member of none.) Every Tg must be a limit point of some S_ρ (because $S = \bigcup_\rho \overline{S_\rho}$), and (in principle) we can choose

$$g^i = \{g^i(n)\}_{n=1,2,\ldots} \to g$$

as $i \to \infty$ such that $Tg^i \in S_\rho$. This will yield Tg via

$$Tg = \lim_{i \to \infty} Tg^i.$$

For example, if $g(n) = n - 1$ we may find an increasing convex sequence converging to g and use case (*ii*). Alternatively we may find an increasing concave sequence converging to g and use case (*iii*).

For which ρ is S_ρ empty? We have no general answer to this; but such ρ certainly exist, for example

(14) $$\rho = (\rho(2), \rho(3), \ldots) = (1, 1, 1, 2, 2, \ldots).$$

If S_ρ were not empty for (14) we should have the obviously incompatible inequalities

(15) $$\begin{cases} f_1 + f_3 < f_2 + f_2, \\ f_2 + f_3 < f_1 + f_4, \\ f_2 + f_4 < f_3 + f_3. \end{cases}$$

Similarly any sequence which commences like

(16) $$\rho = (1, 1, 2, 1, 3, \ldots)$$

is not allowed, and it is easy to produce plenty of other forbidden sequences.

Hammersley [2] proved an asymptotic result for any solution φ of inequality (3) whenever g belonged to a certain subset of S. Namely, if g is non-decreasing for all $n \geqslant \xi > 0$, then the finite convergence of $\sum g(n)/n^2$ is equivalent to the existence of the limit

$$l = \lim_{n \to \infty} \varphi(n)/n,$$

where $-\infty \leqslant l < \infty$. In particular this holds for the maximal solution $f = Tg$ of inequality (3). For example, Harding's case of

$$g(n) = \log \tfrac{1}{2}(n-1) \quad (n > 2)$$

is clearly non-decreasing, and $\sum g(n)/n^2 < \infty$. So

$$l = \lim_{n \to \infty} f(n)/n$$

exists and may be easily calculated. By case (*iii*), $f \in S_\rho$, where ρ is given by equation (2), and iterated solution of equation (6) yields

(17) $$2^{-n} f(2^n) = 2^{-1} g(2) + 2^{-2} g(2^2) + \ldots + 2^{-n} g(2^n).$$

So

(18) $$l = \lim_{n\to\infty} 2^{-n} f(2^n) = -\log 2 + \sum_{i=2}^{\infty} 2^{-i} \log(2^i - 1)$$
$$= 0.253\ldots.$$

4.2.2 PROOF OF RESULTS

Proof of Case (i)

The proof is by induction on n, the result being trivially true for $n = 2$. Suppose it holds for $m = 1, 2, \ldots, n-1$. Iteration of

(19) $$f(m) = f(1) + f(m-1) + g(m) \quad (m < n)$$

yields

(20) $$f(m) = \sum_{i=1}^{m} g(i).$$

So

(21) $$f(n) = f[\rho(n)] + f[n-\rho(n)] + g(n)$$
$$= \sum_{i=1}^{\rho(n)} g(i) + \sum_{i=1}^{n-\rho(n)} g(i) + g(n)$$
$$\geqslant \sum_{i=1}^{n} g(i)$$

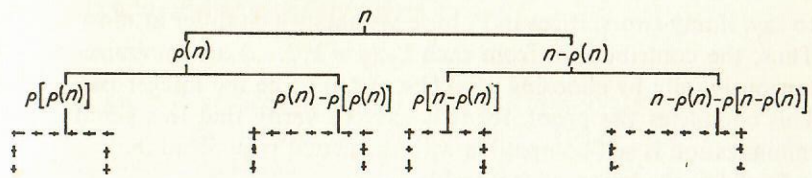
with equality if and only if $\rho(n) = 1$, since g is strictly decreasing. This completes the proof.

Notation for Cases (ii) and (iii)

The iterative decompositions of

(22) $$f(n) = f[\rho(n)] + f[n-\rho(n)] + g(n)$$

may be represented diagrammatically by a tree with a simple *root* vertex with *weight n*, giving birth in the first generation to two vertices weighted $\rho(n)$, $n-\rho(n)$, each of which can create similarly two next generation vertices with suitably representative weights. A vertex fails to divide if it is weighted one. To the two vertices born of the same vertex, with weight w say, we assign the smaller weight $\rho(w)$ to the left-hand vertex of the pair and the larger weight $w-\rho(w)$ to the right-hand vertex. Clearly the weight of vertex A is precisely the number of terminal vertices generated by A and its descendants.



(23)

The solution $f = Tg$ is given by

(24) $$f(n) = \sum_{i\in I} g(a_i),$$

where the summation includes each vertex in the diagram (23), the weight of the ith vertex being a_i.

Vertex A is an *ancestor* of vertex B if B belongs to the subtree of which A is the root. The *depth* of A is the number of ancestors which it possesses, and the *depth* of a subtree is the depth of its root. A tree is *plenary* if the weights of any two vertices of equal depth are the same.

Proof of Case (ii)

We may extend the diagram (23) by allowing any vertex of zero or unit weight to give birth to vertices weighted (zero, zero) or (zero, one) respectively. We set $g(0) = 0$. This does not affect the strict convexity of g, since g is increasing and $g(1) = 0$. Let V_j be the set of depth-j vertices in the extended tree. For each j, V_j has 2^j members with total weight n. Their contribution of

$$\sum_{i=1}^{2^i} g(a_i)$$

to equation (24), where a_i is the weight of the ith vertex, is not changed by the extension. Furthermore we may extend the definition of g from a strictly convex increasing function on the non-negative integers to a convex non-decreasing function on the reals by

(25) $g(x) = (1-x+i)g(i+1) + (x-i)g(i) \quad (x \in (i, i+1)$ for some integer $i)$.

Then

(26) $$\sum_{i=1}^{2^i} g(a_i) \geqslant 2^j g\Big(2^{-j} \sum_{i=1}^{2^j} a_i\Big) = 2^j g(2^{-j} n)$$

with equality if and only if $a_i = 2^{-j} n \ (i = 1, 2, \ldots, 2^j)$ whenever this is integer valued, or otherwise all the a_i lie in the interval of unit length with integer end-points embracing $2^{-j} n$. This is only possible with integer solutions if the spread of the weights of V_j is as small as possible; that is to say, if any two vertices in V_j have weights which differ at most by one. Thus, the contributions from each V_j $(j = 1, 2, \ldots)$ are minimized simultaneously only by choosing $\rho(m)$ $(1 < m \leqslant n)$ to be the integer part of $\frac{1}{2}m$. This completes the proof, for it is easy to verify that this simultaneous minimization is self-compatible with the actual minimization.

Explicit calculation of $f(n)$ yields

(27) $$f(n) = \sum_{i=0}^{[\log_2 n]} 2^i (g([2^{-i}n]) + \{2^{-i}n\} \Delta g([2^{-i}n])),$$

where [ ] and { }, are the integer and fractional parts of their content and $\Delta g(x) = g(x+1) - g(x)$.

Proof of Case (iii)

We order the vertices of a tree as follows: a vertex α is *earlier* than a vertex β (written $\alpha < \beta$) when the depth of α is less than the depth of β, or else when α and β are of equal depth and α lies to the right of β in the same row of the tree. We shall first prove the lemma:

Lemma. *When the vertices of an optimally weighted tree are taken in increasing order of lateness, the weights of the vertices form a non-increasing sequence.*

Proof. Suppose the lemma is not true; and let β_0 be the earliest vertex which is heavier than some still earlier vertex, and let α_0 be the earliest vertex which is lighter than β_0. Writing $w(\alpha)$ for the weight of the vertex α, we can then assert

(28) $$w(\alpha) \geqslant w(\beta) \quad \text{whenever } \alpha \leqslant \beta < \beta_0,$$

while

(29) $$w(\alpha_0) < w(\beta_0), \quad \alpha_0 < \beta_0.$$

Any vertex is both lighter and later than any of its ancestors; so the inequality (29) ensures that α_0 cannot be an ancestor of β_0 nor vice versa. Therefore, α_0 and β_0 have a latest common ancestor γ distinct from both. Let $\alpha_1, \alpha_2, \ldots, \alpha_p$ be the successive ancestors of α_0 leading up to $\alpha_p = \gamma$; and let $\beta_1, \beta_2, \ldots, \beta_q$ be corresponding ancestors of β_0 leading up to $\beta_q = \gamma$. Since $\alpha_0 < \beta_0$, we see that α_0 is not deeper than β_0, which implies

(30) $$p \leqslant q.$$

If $q = 1$, then $p = 1$, and γ is the immediate common ancestor of both α_0 and β_0, which implies

(31) $$w(\beta_0) = \rho[w(\gamma)] \leqslant w(\gamma) - \rho[w(\gamma)] = w(\alpha_0)$$

in contradiction of the inequality (29). Hence

(32) $$q \geqslant 2.$$

If α and β are two vertices satisfying $\alpha \leqslant \beta$, the immediate ancestor of α is not later than the immediate ancestor of β. Hence $\alpha_0 < \beta_0$ implies $\alpha_1 \leqslant \beta_1 < \beta_0$, which implies successively

(33) $$\alpha_i \leqslant \beta_i < \beta_{i-1} < \ldots < \beta_0 \quad (i = 1, 2, \ldots, p).$$

Writing

(34) $$a_i = w(\alpha_i), \quad b_i = w(\beta_i),$$

we obtain from the inequalities (28), (29) and (33)

(35) $$a_0 < b_0, \quad a_i \geqslant b_i \quad (i = 1, 2, \ldots, p).$$

Now consider the effect of interchanging the two subtrees with roots at α_0 and β_0. Most of the weights in the tree will merely be shifted in position without changing their magnitudes; and the only weights to suffer a change of magnitude will be those at $\alpha_1, \alpha_2, \ldots, \alpha_{p-1}$ and $\beta_1, \beta_2, \ldots, \beta_{q-1}$. The former will be changed from $a_1, a_2, \ldots, a_{p-1}$ to $a_1 + \delta, a_2 + \delta, \ldots, a_{p-1} + \delta$, and the latter from $b_1, b_2, \ldots, b_{q-1}$ to $b_1 - \delta, b_2 - \delta, \ldots, b_{q-1} - \delta$, where

(36) $$\delta = b_0 - a_0 > 0.$$

Thus the total score (24) for the new tree will exceed the total score for the original tree by an amount

(37) $$\Delta = \sum_{i=1}^{p-1} [g(a_i + \delta) - g(a_i) + g(b_i - \delta) - g(b_i)] + \sum_{i=p}^{q-1} [g(b_i - \delta) - g(b_i)].$$

In equation (37), we interpret a void sum

$$\left(\text{such as } \sum_{i=1}^{0} \text{ or } \sum_{i=p}^{p-1}\right)$$

as zero if it should occur. However, the first sum on the right of equation (37) can only be void if $p = 1$, in which case the inequality (32) prevents the second sum from being void. So at least one of the sums in equation (37) is non-void, and we shall be able to conclude that

(38) $$\Delta < 0$$

provided that each individual term actually occurring on the right of equation (37) is strictly negative. And this is true; because the inequality (36) implies

(39) $$g(b_i - \delta) - g(b_i) < 0$$

inasmuch as g is a strictly increasing function. Also, since

$$b_i-\delta<b_i\leqslant a_i<a_i+\delta,$$

the chord of the strictly concave function g from $b_i-\delta$ to $a_i+\delta$ must lie strictly below the chord from b_i to a_i, and therefore the mid-point of the former chord lies strictly below the mid-point of the latter chord. Hence

(40) $$g(a_i+\delta)+g(b_i-\delta)-g(a_i)-g(b_i)<0.$$

This establishes the inequality (38). However (38) contradicts the assumption that the original tree was optimally weighted in the sense of having a minimum total score; and this contradiction proves the lemma.

Now consider an optimally weighted tree. Descend from its root, always taking the left-hand branch, until a vertex V of unit weight is reached (at depth D, say). Let v be any other vertex of depth D. Then $v<V$, and the lemma gives $w(v)\geqslant 1$. Suppose $w(v)>2$. Then v has a right-hand descendant v' at depth $D+1$ such that $v'>V$ and

$$w(v')\geqslant \tfrac{1}{2}w(v)>1=w(V),$$

contradicting the lemma. So $1\leqslant w(v)\leqslant 2$. The vertices at depth D, read as a row from right to left, have non-increasing weights. Hence, if the weights in the right-hand half of the row are not all 2, the weights in the left-hand half must all be 1. In either case, one of the two depth-1 subtrees of the original tree will have equal weights of unity in its last row, and so will be plenary. The weight of its root will be a power of 2, say 2^i. The weight of the other depth-1 subtree must be at least 2^{i-1} and at most 2^{i+1}. If the left-hand subtree is plenary, we get $2.2^i\leqslant n\leqslant 3.2^i$; and if the right-hand subtree is plenary, we get $\tfrac{3}{2}.\ 2^i\leqslant n\leqslant 2.2^i$. In either case, $\tfrac{1}{3}n\leqslant 2^i\leqslant \tfrac{2}{3}n$ holds. This completes the proof of equation (2).

Explicit calculation of $f(n)$ now yields

(41) $$f(n)=\sum_{k=0}^{i}(2^k-1)g(2^{i-k})+\left[\frac{n-2^i}{2^i-k}\right](g(2^{i-k+1})-g(2^{i-k}))$$
$$+g\left(2^{i-k}\left(1+\left\{\frac{n-2^i}{2^{i-k}}\right\}\right)\right),$$

where $1<2^{-i}n\leqslant 2$, and [ ] and { } indicate the integer and fractional parts of their content.

4.2.3 UNLABELLED SHAPES

In connection with his *unlabelled shapes*, Harding considered a modified form of equation (1) with the factor 2 in the numerator $2M(r)M(n-r)$ suppressed whenever $r=n-r$. This leads to a very irregular sequence ρ,

which he tabulated for $n\leqslant 160$. We have made some fragmentary but very limited progress with this more difficult problem, which in our notation reads as

(42) $$f(1)=0,\quad f(n)=\min_{1\leqslant r\leqslant \frac{1}{2}n}[f(r)+f(n-r)+g(n)+c\delta_{r,n-r}],$$

where

(43) $$g(n)=\log\tfrac{1}{2}(n-1),\quad c=\log 2,$$

and δ is the Kronecker delta.

In the first place, this new f is a solution of inequality (5) with g replaced by $g(n)+c=\log(n-1)$; and, since $\sum n^{-2}\log(n-1)$ converges, we still have $f(n)/n$ tending to a limit as $n\to\infty$. We have made computer calculations of f for $n\leqslant 8000$; and the limit appears to be approximately 0.587.

Secondly, in place of the previous relation $\rho(n)\geqslant n/3$, we now have

(44) $$\rho(n)>n/5.$$

To see this, consider the tree (23); and for brevity write

(45) $$\alpha=\rho(n),\quad \beta=\rho[n-\rho(n)],\quad \gamma=n-\rho(n)-\rho[n-\rho(n)]$$

for the initial splitting of n. The definition of ρ ensures that

(46) $$\beta\leqslant\gamma;$$

and, of course, we have $\alpha+\beta+\gamma=n$. We consider interchanges of subtrees, as in Section 2; and we say that the *penalty* c is *incurred* wherever a non-zero Kronecker delta occurs.

Suppose that the tree is optimally weighted. Then we cannot have

(47) $$\alpha<\beta;$$

for in this event an interchange of α and β cannot incur a penalty at the lower level because $\alpha<\gamma$ in view of inequalities (46) and (47), nor at the higher level because $\beta\leqslant\tfrac{1}{2}[n-\rho(n)]<\tfrac{1}{2}n$. The change of score must be non-negative, since the tree was optimally weighted. Thus

$$g(\alpha+\gamma)-g(\beta+\gamma)\geqslant 0,$$

contradicting inequality (47). If $\alpha=\beta$, $\gamma=\tfrac{1}{2}n$, then $\alpha=\tfrac{1}{4}n$ and inequality (44) is trivial. So we consider

(48) $$\text{either}\quad \alpha>\beta\quad\text{or}\quad \gamma\neq\tfrac{1}{2}n,\quad \alpha\geqslant\beta.$$

An interchange of α and γ can only incur one penalty at most, perhaps at the upper level in the former case of (48), and perhaps at the lower level in the latter case of (48). The change of score is at most

(49) $$g(\alpha+\beta)+\log 2-g(\beta+\gamma)\geqslant 0,$$

whence

(50) $$\beta+\gamma-1\leqslant 2(\alpha+\beta-1)\leqslant 4\alpha-2,$$

by (48). Hence $n = \alpha+\beta+\gamma \leqslant 5\alpha - 1$; and $\rho(n) = \alpha > n/5$.

Finally, we show that the penalty is never incurred for $n > 2$. Suppose that the penalty is incurred for $n = 2m$, with $m > 1$ and m as small as possible. Then

(51) $$2f(m) = 2f(\tfrac{1}{2}n) \leqslant f(r) + f(n-r) - \log 2 \quad (1 \leqslant r < m).$$

With $r = m-1$, this yields

(52) $$\Delta^2 f(m-1) \geqslant \log 2,$$

where $\Delta^2 h$ is the second forward difference of a function h, and (later) Δh is the first forward difference. For $2 < s \leqslant m+1 < 2m$ the penalty is not incurred, by definition of m, and hence

(53) $$f(s) = f[\rho(s)] + f[s-\rho(s)] + g(s),$$

(54) $$f(s+1) \leqslant f[\rho(s)+1] + f[s-\rho(s)] + g(s+1),$$

whence

(55) $$\Delta f(s) \leqslant \Delta f[\rho(s)] + \Delta g(s).$$

Write

(56) $$r_j = m, \quad r_{j-1} = \rho(r_j), \ldots, \quad r_0 = \rho(r_1),$$

where j is to be chosen presently. Substituting $s = r_i \,(1 \leqslant i \leqslant j)$ in inequality (55) and adding, we get

(57) $$\Delta f(m) \leqslant \Delta f(r_0) + \sum_{i=1}^{j} \Delta g(r_i).$$

We can write inequality (44) as

(58) $$\tfrac{1}{5}n < \rho(n) \leqslant \tfrac{1}{2}n;$$

and so equations (56) and (58) provide

(59) $$r_i \geqslant 2^i r_0.$$

However

(60) $$\Delta g(n) = \log\left(1 + \frac{1}{n-1}\right),$$

which is a decreasing function of n. So

(61) $$\sum_{i=1}^{j} \Delta g(r_i) \leqslant \sum_{i=1}^{j} \log\left(1 + \frac{1}{2^i r_0 - 1}\right) \leqslant \sum_{i=1}^{\infty} \frac{1}{2^i r_0 - 1}$$

$$\leqslant \frac{2r_0}{2r_0 - 1} \sum_{i=1}^{\infty} \frac{1}{2^i r_0} = \frac{2}{2r_0 - 1}.$$

Hence

(62) $$\Delta f(m) \leqslant \Delta f(r_0) + \frac{2}{2r_0 - 1}$$

Similarly, instead of equations (53) and (54), we have

(63) $$f(s+1) = f[\rho(s+1)] + f[s+1-\rho(s+1)] + g(s+1),$$

(64) $$f(s) \leqslant f[\rho(s+1)-1] + f[s+1-\rho(s+1)] + g(s),$$

which yield

(65) $$\Delta f(s) \geqslant \Delta f[\rho(s+1)-1] + \Delta g(s) \geqslant \Delta f[\rho(s+1)-1].$$

This time we write, with J to be chosen presently,

(66) $$R_J = m-1, \quad R_{J-1} = \rho(R_J+1)-1, \ldots, \quad R_0 = \rho(R_1+1)-1.$$

Substituting $s = R_i \,(1 \leqslant i \leqslant J)$ in inequality (65) and adding, we have

(67) $$\Delta f(m-1) \geqslant \Delta f(R_0).$$

From inequalities (52), (62) and (67), we deduce

(68) $$\log 2 \leqslant \frac{2}{2r_0 - 1} + \Delta f(r_0) - \Delta f(R_0).$$

From equations (56), (58) and (66), we see that

(69) $$\tfrac{1}{5}r_i < r_{i-1} \leqslant \tfrac{1}{2}r_i, \quad \tfrac{1}{5}(R_i+1) < R_{i-1}+1 \leqslant \tfrac{1}{2}(R_i+1).$$

Hence, for $m \geqslant 174$, we can always choose j and J such that

(70) $$6 \leqslant r_0 \leqslant 30, \quad 6 \leqslant R_0 \leqslant 34.$$

Now inequalities (68) and (70) give the contradiction

(71) $$0.693 < \log 2 \leqslant \tfrac{2}{11} + \max_{6 \leqslant r \leqslant 30} \Delta f(r) - \min_{6 \leqslant R \leqslant 34} \Delta f(R) < 0.623.$$

Here we have used the computed values of $\Delta f(n)$ for $6 \leqslant n \leqslant 34$. Thus the first incurrence (after $n = 2$) of a penalty cannot occur for $n = 2m \geqslant 348$. However, the computer studies also show that no penalty is incurred for $2 < n < 348$. This completes the proof that the penalty is never incurred for $n > 2$.

It is possible to have $\rho(n) < \tfrac{1}{3}n$; and it seems likely from the computer studies that this happens infinitely often. At least is happens for the following values of n:

Table 1. Values of $n \leqslant 8000$ for which $\rho(n) < \frac{1}{3}n$

n	$n-3\rho(n)$	n	$n-3\rho(n)$	n	$n-3\rho(n)$	n	$n-3\rho(n)$
4	1	1369	1	2758	4	5444	14
10	1	1370	2	5392	7	5456	2
22	1	1372	4	5400	3	5461	7
168	3	1374	6	5408	2	5466	12
343	1	1388	2	5414	8	5492	2
345	3	2700	6	5418	6	5507	5
669	6	2702	8	5435	5	5508	6
671	8	2714	2	5437	7	5524	4
682	1	2719	7	5438	8	5525	5
683	2	2722	10	5439	9	5529	9
1346	2	2744	2	5440	10		
1352	8	2757	3	5442	12		

We have attempted to measure the *popularity* $P(n)$ of an integer n by counting the number of integers i for which either $n = \rho(i)$ or $n = i-\rho(i)$ and assigning this number to $P(n)$. Then $P = \{P(n)\}_{n=2,3,\ldots}$ is an infinite sequence of integers which we have calculated for $n \leqslant 2666$. The popularity sequence P_1 for solution (2) to equation (1) is

$$P_1 = \{P_1(n)\}_{n=2,3,\ldots} = (5, 2, 8, 2, 2, 2, 14, 2, \ldots) \tag{72}$$

or

$$P_1(n) = \begin{cases} 2+2^{i-1}+2^i & \text{if } n = 2^i, \\ 2 & \text{otherwise,} \end{cases} \tag{73}$$

and indicates the form of the most popular integers. Continuing this analogy with equation (2), we conjectured that in order to find the 'best' r ('best' by some unknown rule) to minimize the right-hand side of equation (42) it is only necessary to scan the numbers $\{f(r)\}_{r=1,\ldots,n-1}$ without consideration of the complementary set $\{f(n-r)\}_{r=1,\ldots,n-1}$. (For example equation (2) defines the 'best' r for Harding's *labelled shapes* as a power of 2 lying in $[\frac{1}{3}n, \frac{2}{3}n]$.) Thus, in order to be consistent with our notion of popularity, we desired that for each n the chance choice of $n-r$ should have no effect upon P. So from P we constructed an amended sequence P^* of popularities by taking one away from the lesser of $P[\rho(n)]$ and $P[n-\rho(n)]$ for each $n \leqslant 8000$. (The amended sequence P_1^* for equation (7) is

$$P_1^* = \{P_1^*(n)\}_{n=2,3,\ldots} = (4, 0, 7, 0, 0, 0, 13, 0, \ldots) \tag{74}$$

or

$$P_1^*(n) = \begin{cases} P_1(2^i)-1 & \text{if } n = 2^i, \\ 0 & \text{otherwise,} \end{cases} \tag{75}$$

which exhibits clearly the nature of equation (2).) Our sequence P^* for unlabelled shapes has non-zero entries as given in Table 2.

Table 2. Values of $n \leqslant 2666$ for which $P^*(n) > 0$

n	$P^*(n)$	n	$P^*(n)$	n	$P^*(n)$	n	$P^*(n)$	n	$P^*(n)$
3	4	219	2	487	1	1133	1	1834	474
4	1	220	3	568	2	1141	6	1838	28
5	2	221	36	664	1	1158	1	1839	28
7	†6	227	112	675	1	1169	2	1840	121
8	†6	229	99	677	3	1207	2	1843	14
12	†2	231	10	685	3	1247	1	1844	77
15	16	232	†1	696	11	1253	2	1845	13
17	†2	233	23	751	2	1265	1	1848	19
20	5	235	†4	783	4	1341	4	1849	55
27	11	236	11	809	2	1360	9	1853	†11
28	3	240	25	885	14	1391	10	1856	4
31	1	245	†4	888	†4	1393	2	1857	19
32	9	247	1	889	†4	1589	2	1858	1
35	14	253	2	891	7	1612	1	1861	9
42	5	255	†4	895	14	1614	8	1865	8
47	1	295	2	898	60	1701	1	1866	31
52	2	323	†2	904	317	1783	8	1944	7
55	26	335	4	906	106	1787	3	2035	1
58	2	341	2	911	65	1789	8	2059	14
59	15	353	2	913	50	1793	7	2076	2
60	4	355	1	914	114	1795	16	2171	14
62	3	408	†1	916	111	1799	29	2258	1
67	16	430	†2	917	1	1802	246	2264	3
77	5	437	8	918	261	1803	11	2266	1
90	8	441	†2	921	5	1804	33	2278	14
97	7	443	1	922	40	1807	23	2322	1
102	3	448	97	925	1	1810	370	2326	1
107	11	450	27	926	40	1815	153	2483	1
113	19	456	331	927	4	1817	39	2485	1
114	77	457	1	930	20	1818	176	2507	2
115	17	458	6	931	26	1819	30	2619	14
117	4	460	29	935	30	1820	39		
119	5	462	62	937	1	1822	26		
121	†4	464	7	942	24	1824	4		
122	3	465	10	944	1	1825	108		
126	12	468	16	948	†11	1827	17		
127	2	469	19	949	1	1829	11		
129	2	473	28	980	3	1830	240		
181	6	475	9	1026	4	1831	24		
216	†4	476	7	1131	1	1832	149		

† P^* was first constructed for values of n such that $P[\rho(n)] \neq P[n-\rho(n)]$. For integers excluded by this condition the comparison process was repeated using P^* as the yardstick. However, for the pairs $[\rho(n), n-\rho(n)] = (7, 8)$, $(12, 17)$, $(121, 216)$, $(232, 408)$, $(235, 255)$, $(245, 255)$, $(323, 441)$, $(430, 441)$, $(888, 889)$, $(948, 1853)$ both $P[\rho(n)] = P[n-\rho(n)]$ and $P^*[\rho(n)] = P^*[n-\rho(n)]$. These values are marked by a dagger.

Table 2 shows that the most popular integers are distributed about the numbers 15, 27, 55, 114, 228, 456, 912, 1824. We do not know why these numbers are chosen, nor have we made any significant progress towards understanding their associated distributions. Indeed, further investigation seems likely to be challenging since any exhaustive quantitative analysis will be very involved. More details of $f(n)$, $\rho(n)$, $P(n)$ $(n \leqslant 8000)$ may be obtained from us.

We have just heard of a paper [3], which appears (we have so far only seen the summary and not the paper itself) to deal with closely related, if not identical, material.

REFERENCES

1. E. F. Harding, "The probabilities of rooted tree-shapes generated by random bifurcation", *Adv. Appl. Prob.* 3 (1971), 44–77. (See also 4.1 of this book.)
2. J. M. Hammersley, "Generalization of the fundamental theorem on sub-additive functions", *Proc. Camb. Phil. Soc.* **58** (1962), 235–238.
3. M. Fredman and D. Knuth, "Recurrence relations based on minimization", *Stanford University Computer Science Report No.* 248 (1972).

5

NAVIGATION IN THE PRESENCE OF AN UNCERTAINTY