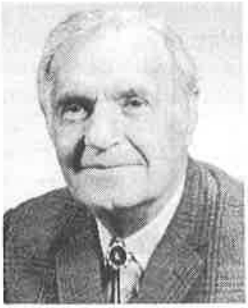
PROBLEM PROPOSING AND MATHEMATICAL CREATIVITY

(Reprinted from Crux Mathematicorum, Vol 12, No 10, Dec 1986)

M.S. Klamkin



Professor Murray S. Klamkin is one of the world's greatest problem creators and solvers. He was the dedicated coach of the USA IMO teams from the 19th (1978) through the 26h (1985) IMO. He was editor of that excellent Olympiad Corner in Crux Mathematicorum (1979-1985) and has recently produced the MAA book "International Mathematical Olympiads 1979-1985 and Forty Supplementary Problems".

There are very many mathematical competitions, at many levels, being given around the world each year. For a treatment of a number of these, see the Proceedings of the congress ICME 5 and the bibliography in [1]. As a consequence of these very many competitions, there have been quite a number of duplications of the problems, either inadvertently or in some cases by direct copying. Also, many contest problems have appeared previously in well known books and journals. Since these competitions have apparently become increasingly more important, there are training sessions which help prepare for a number of them. In view of all this, competition examination committees now have to be much more vigilant than ever before in setting their competitions. They must now continually keep abreast of problems set in other competitions. They have to be very careful about duplicating problems from books and journal problem sections. Even if the book or journal used is not too well known, the problems used could have already been duplicated in other books and journals that are well known. To play safe, problems should either be new or else based on some nice result from some non-recent mathematical paper. This brings us back to the theme of this congress session, "How does one create new problems?"

In almost all my previous papers concerned with problem solving and proposing, I have highly recommended the following five books of George Polya. To me they are still the best books around dealing with the subject.

How to Solve It, Doubleday, New York, 1957.
Mathematical Discovery, Vol. I, Wiley, New York, 1962.
Mathematical Discovery, Vol II, Wiley, New York, 1965.
Induction and Analogy in Mathematics, Princeton University Press, Princeton, 1954.
Patterns of Plausible Inference, Princeton University Press, Princeton, 1954.

In this paper, I will be concerned with the creative aspects of problem solving and proposing. Although the psychological aspects of creativity in mathematics are important, I will dwell mostly on the mathematical aspects. This is mainly due to the fact that I do not know much about the psychological aspects.

In solving or creating problems it is certainly quite helpful to have a good memory and to be observant. George Polya makes the analogy of finding a precious uncut stone on the shore and tossing it away since it is not recognized as being valuable. One has to do a certain amount of cutting and polishing before the value of the stone is recognized, although an expert usually can get away with just a careful examination. So in regard to a problem which has just been solved or whose solution has been looked up, we should not immediately pass on to something else. Rather, we should "stand back" and re-examine the problem in light of its solution and ask ourselves whether or not the solution really gets to the "heart" of the problem. Mathematically, one of the points being made here is to check whether or not the hypotheses of the problem are necessary for the result. (That the hypotheses are sufficient follows from the validity of the result). Additionally, although our solution may be valid, there may be and usually are better ways of looking at the problem which make the result and the proof more transparent and can as well lead to extensions. Consequently, it should be easier to understand the result and the proof as well as to give a non-trivial extension. I will illustrate these remarks by considering a number of elementary mathematical results and will show how, by careful re-examination, one can be led to more general results, some of which are considerably more sophisticated. How well one will succeed in this process will of course depend on one's powers of observation, knowledge, memory, and persistence, in addition to any natural creative ability.

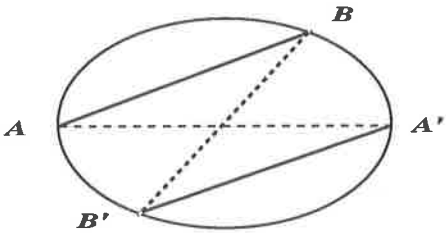
1. Chords and Diagonals

For the first illustration, let us consider the intuitive and known result that the largest chord of a circle is a diameter. One simple proof follows by connecting the end points A, B of the chord to the center O of the circle and using the triangle inequality:

$$AB < AO + BO = \text{diameter}.$$

1.1 Usually, many students and teachers do not bother with the proof, and those that do usually pass on quickly to something else since the proof is so simple. However, as mentioned previously, let us "stand back" and examine the result in light of its proof. One of the questions to ask ourselves is "are all the properties of the circle necessary for this result?"

One possibility for a generalization is to enlarge the class of figures from circles to those which are centrosymmetric. Again it is intuitive (but perhaps not as much as before) that the largest chord must contain the center. The following proof is just a slight modification of the previous one. Referring to the figure, we see that



$$AB < AO + BO \le 2 \max \{AO, BO\} = \max \{AA', BB'\}$$

It is to be noted that this last result is valid for any centrosymmetric set of points in any dimension. Basically, what has been proved is that the longest chord of a parallelogram is the longest diagonal.

1.2 Again there may be a temptation to either rest or pass on to something else. But let us be persistent and see if we cannot extend the last result. Naturally, how we succeed here is a function of our persistence, our knowledge, our experience, etc. Usually, there will be many false starts and lots of wasted paper, but that is all part of the process.

The parallelogram idea leads one to consider the analogous problem of determining the longest chord of a polygon or even a polytope in any dimension. We will show that the endpoints of the longest chord are two of the vertices of the polygon or polytope. This is also a known result. A geometric proof for polygons is given by Rademacher and Toeplitz [2] and a vector proof for polytopes is given in [3]. For completeness here, we give the vector proof.

We need only consider convex polytopes. For if the result is valid for the convex hull of an arbitrary polytope, it is also valid for the polytope.

Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ denote vectors from a common origin to the vertices of the polytope. Then

$$\vec{r}\,' = \Sigma\, w_1' \vec{v}_1, \qquad \text{and} \qquad \vec{r} = \Sigma\, w_j \vec{v}_j$$

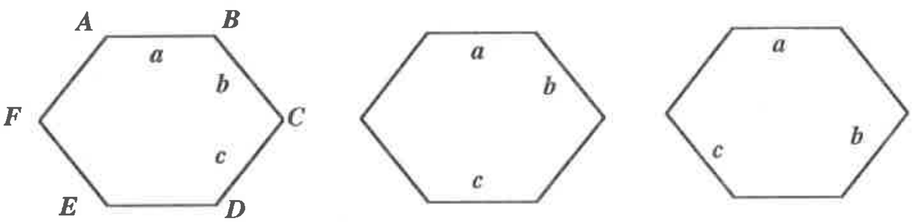
where $w_i'.w_j \ge 0$ and $\Sigma w_i' = \Sigma w_j = 1$, will denote two vectors from the common origin to two points within or on the boundary of the polytope. Using the triangle inequality repeatedly and the properties of w_i'. and w_j, we get the following sequence of inequalities:

$$|\vec{r}\,' - \vec{r}| = |\sum_i w_i' (\vec{v}_i - \vec{r})| \le \sum_i w_i' |(\vec{v}_i - \vec{r})| \le \operatorname{Max}_i |\vec{v}_i - \vec{r}|$$

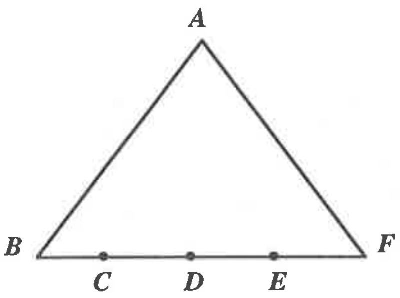
$$\leq \underset{i}{\text{Max}} \left| \sum_j w_j (\vec{v}_i - \vec{v}_j) \right| \leq \underset{i}{\text{Max}} \sum_j w_j |\vec{v}_i - \vec{v}_j| \leq \underset{i,j}{\text{Max}} |\vec{v}_i - \vec{v}_j| \quad .$$

1.3 At this stage, we may again be tempted to pass on to something else. However, we will still persist in looking for other extensions. Here is where a knowledge base and memory comes in. After a while, triggered off by the above results, I remembered the result that for any convex quadrilateral there is a least one side which is smaller than the greater of the diagonals of the quadrilateral [4]. Then after playing around (experimenting), I conjectured that for any convex n-gon, there are at least $n - 2$ sides which are shorter than the longest diagonal, and furthermore, $n - 2$ is best possible. After more playing around, I came up with the following proof. For simplicity, I illustrate the proof for a hexagon; however, the proof carries through for an n-gon.

Our proof is indirect. We assume that there are at least three sides of the hexagon which are greater than or equal to the longest diagonal. Labelling these three sides as a, b, c we have the following possible configurations:



The "worst" configuration for our proof will be the first one. Since in a triangle the greatest angle is opposite the greatest side, we have: in ΔABC, $\angle B \leq \frac{\pi}{3}$; in ΔBCD, $\angle C \leq \frac{\pi}{3}$; in ΔCDE, $\angle D \leq \frac{\pi}{2}$; in ΔDEF, $\angle E \leq \pi$; in ΔEFA, $\angle F \leq \pi$; and in ΔFAB, $\angle A \leq \frac{\pi}{2}$. This leads to a contradiction since the sum of the interior angles of an n-gon is $(n - 2)$. For the other two configurations, the sum of the interior angles would be bounded by a number even smaller than $\frac{11\pi}{3}$. The next figure shows that $n - 2$ is best possible (for $n = 6$).



$\angle BAF < \frac{\pi}{3}$, $AB = AF$

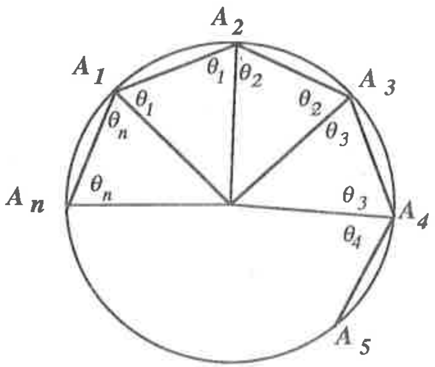
No doubt a person with more persistence, knowledge and creativity can find further extensions of the previous results. The subsequent illustrations will be treated in a less extended way.

2. Equilateral and Equiangular Polygons [5].

It is a well known elementary result that all equilateral triangles are equiangular and, conversely, that all equiangular triangles are equilateral. One simple proof follows by considering the equilateral or equiangular triangle to be congruent to itself in different ways. Since the proof is nice and simple and the result does not extend to polygons, we usually pass on to something else. Note that a rhombus need not be a square and a rectangle need not be a square.

2.1 Before quitting the search for an extension, we should check whether or not we have inadvertently left out some conditions. One difference between triangles and higher order polygons is that a triangle always has a circumcircle. So let us restrict the class of polygons to those which can be inscribed in a circle. Now it follows easily that all inscribed equilateral polygons are equiangular but not conversely (which again follows by considering the rectangle).

2.2 Let us re-examine why the rectangle is a counter example. Referring to the figure, where it is assumed the n-gon is equiangular, we obtain the simultaneous set of linear equations



$$\begin{aligned}\theta_1 + \theta_2 &= k,\\ \theta_2 + \theta_3 &= k,\\ &\vdots\\ \theta_n + \theta_1 &= k,\end{aligned}$$

from which

$$\begin{aligned}\theta_1 &= \theta_3 = \theta_5 = \ldots = \theta_{n-1},\\ \theta_2 &= \theta_4 = \theta_6 = \ldots = \theta_n .\end{aligned}$$

If n is odd, the angles are constant and the polygon is regular; if n is even, the polygon needs not be regular. This latter result is ascribed by H.S.M. Coxeter to M. Riesz.

2.3 The figure and the set of equations suggest an extension of the result of Riesz. We first define a d-diagonal of an n-gon, with $2d \leq n$, as a diagonal which "skips" d - 1 vertices (e.g., in the above figure, A_1A_2 is a 1-diagonal and A_1A_4 is a 3-diagonal). One can now show that if all the d-diagonals (d fixed) of an inscribed n-gon are congruent and $(n,d) = 1$, then the polygon is regular (when $d = 2$, we get Riesz's result). For a proof and similar results for circumscribed polygons, see [2].

2.4 Instead of going from a triangle to a polygon, we can consider the 3-dimensional extension to tetrahedra. Here if all the edges are congruent, it follows easily that all the face angles are congruent and that all the dihedral angles are congruent. I leave it as an exercise to show that if all the dihedral angles are congruent then the tetrahedron is regular, and to extend this result to n-dimensional simplexes, for which there are many different sets of angles to consider.

3. Intersecting Curves on a Surface [6].

It is well known that two great circles of a sphere always intersect in two points which are antipodal. The usual proof follows by noting that the planes of the two circles must both contain the center of the sphere and thus the planes have a line of intersection which includes a common diameter of the two great circles.

3.1 Although a first proof of a given result can lead to an extension, in many cases it is an alternate proof that more readily leads to the extension. Another proof which appears more basic is that since each circle divides the surface into two congruent regions, the circles must intersect. This gives the following generalization whose proof is identical except for the replacing of "circles" by "curves".

> *If two simple closed centrosymmetric curves lie on a simple closed centrosymmetric surface homeomorphic to a sphere, with all three having the same center, then the two curves intersect.*

By centrosymmetry, the points of intersection will occur in pairs of antipodal points. Also, it is to be noted that the result is not valid on a multiconnected surface, e.g. a torus.

3.2 On looking back at the previous proof it will be seen that the condition of centrosymmetry is not necessary. It was essential only that the surface and the two curves each be mapped into themselves by a reflection through a common point. We now formalize these ideas. Let 0 denote an interior point of a three-dimensional region starlike with respect to 0, and with boundary S, and consider a mapping f of S into itself such that each point of S goes into its antipodal point with respect to 0. The result in 3.1 can now be extended to:

> *If C_1, and C_2, are two simple closed curves on S such that C_1 and C_2 each map into themselves under the mapping f, then C_1 and C_2 intersect in pairs of antipodal points with respect to 0.*

3.3 On re-examination of this result and its proof, we see that a wider generalization is possible which is given by:

> *If C_1 and C_2 are two simple closed curves on a simple closed surface S and if f is a continuous mapping, without fixed points, of S into itself such that $C_1 = f(C_1)$ and $C_2 = f(C_2)$, then $C_1 \cap C_2 \neq 0$.*

A proof follows by using the Brouwer fixed point theorem; see [6].

Even though what one should be looking for in the re-examination of known results has been indicated in the above examples, re-examination does not automatically lead to further non-trivial results. To aid in this re-examination, we consider some heuristics given by Polya in his books mentioned earlier, namely: specialize, generalize and make analogies. I now give some further illustrations.

4. Maximum Number of Terms in a Sequence

(IMO, 1977, #2) *In a finite sequence of real numbers, the sum of any seven successive terms is negative, and the sum of any eleven successive terms is positive. Determine the maximum number of terms in the sequence.*

Let the sequence be $A_1, A_2, \ldots\ldots,$ By considering the array

$$\begin{array}{cccc} A_1 & A_2 & \cdots & A_7 \\ A_2 & A_3 & \cdots & A_8 \\ A_3 & A_4 & \cdots & A_9 \\ \vdots & \vdots & & \vdots \\ A_{11} & A_{12} & & A_{17} \end{array}$$

it follows that the number of terms is less than 17, and by an explicit construction it can be shown that 16 is the maximum.

4.1 One possible generalization is to vary the numbers given in the problem. This was done by the English contestant John Rickard, who deservedly won a special prize for it. He replaced 7 and 11 by two relatively prime integers p and q and then showed that the maximum number is $p + q - 2$. A further extension appeared in ***Mathematical Spectrum*** 12 (1979/80) 61 where now $(p,q) = d$. A still further extension is to require the additional condition that the sum of any r successive terms is zero.

5. Isoperimetric Inequality [7].

It is known that the isoperimetric quotient I.Q. $= A/P^2$ for triangles of area A and perimeter P is a maximum for the equilateral case.

5.1 On "playing around" with this, one is led to the following:

Given that A_1 is an interior point of an equilateral triangle ABC and A_2 is an interior point of triangle A_1BC, then

$$\text{I. Q.}(A_1BC) > \text{I.Q.}(A_2BC).$$

One expects this inequality since one feels intuitively that A_1BC is "closer" to being an equilateral triangle than A_2BC is.

5.2 By increasing the dimensionality of the problem, one can consider the analogous inequalities for a simplex in E^n, for which there are many different isoperimetric quotients. In particular for E^3, given that A_1 is an interior point of a regular tetrahedron $ABCD$ and that A_2 is an interior point of tetrahedron A_1BCD, then I.Q.(A_1BCD) > I.Q.(A_2BCD) where here the isoperimetric quotient of a tetrahedron T is defined by

$$\text{I.Q.}(T) = \text{Vol}\,(T) / [\text{Area}\,(T)]^{3/2}$$

A proof is given in [4]. For another isoperimetric quotient, replace Area$^{3/2}$ by the total edge length cubed.

6. Maximum Volume of a Tetrahedron.

(IMO, 1967, #2) Prove that if one and only one edge of a tetrahedron is greater than 1, then its volume is $\leq \frac{1}{8}$.

6.1 We can extend the problem simultaneously in two different ways as in [8]:

Determine the maximum volume of an n-dimensional simplex if at most r edges are greater in length than 1 ($r = 1, 2 \ldots., n$).

This problem is still open. The special case $n = 3$, $r = 2$ or 3 is solved in [9]. For the $r = 3$ case it is also assumed that the 3 edges longer than 1 cannot all be concurrent, otherwise the volume can be unbounded.

7. A Two Triangle Inequality

(Putnam, 1982, B-6) If $K(x,y,z)$ denotes the area of a triangle of sides x, y, and z, prove that

$$\sqrt{K(a,b,c)} + \sqrt{K(a',b',c')} \leq \sqrt{K(a + a',\, b + b',\, c + c')}$$

This inequality is a special case of one derived by me in [10], and I now give the motivation for its derivation. A well known elementary inequality for the sides a, b, c of a triangle is

$$abc \geq (a+b-c)(b+c-a)(c+a-b) \qquad (1)$$

with equality if and only if $a = b = c$. For a variety of different proofs of this, and for other references, see {8}. One of the simplest proofs is to square both sides and note that $a^2 \geq a^2 - (b - c)^2$, etc. Apparently this proof does not readily lead to any significant extensions. However, by interpreting (1) geometrically, we are led to several generalizations by an averaging process over the sides of the triangle.

We consider here another triangle $A'B'C'$ where

$$a' = \frac{(b+c)}{2}, \qquad b' = \frac{(c+a)}{2}, \qquad c' = \frac{(a+b)}{2}$$

Since $s = s'$ (equal semi-perimeters), and triangle $A'B'C$ is "closer" to an equilateral triangle than ABC, we should expect that $K(A'B'C') \geq K(ABC)$. Since $8K^2(A'B'C') = abcs$, the latter inequality is equivalent to (1).

More generally, we should expect the same area inequality for any reasonable averaging transformation which makes triangle A'B'C "more equilateral" than ABC. More precisely, if

$$a' = ua + vb + wc,$$
$$b' = va + wb + uc,$$
$$c' = wa + ub + vc,$$

where

$$u + v + w = 1 \qquad u, v, w \geq 0,$$

then $s' = s$ and $K(A'B'C') \geq K(ABC)$. This last triangle inequality is equivalent to

$$(xa+yb+zc)\,(ya+zb+xc)\,(za+xb+yc) \geq (a+b-c)\,(b+c-a)\,(c+a-b)$$

where

$$x + y + z = 1, \qquad -1 \leq x, y, z \leq 1.$$

7.1 We can generalize further by letting a_i, b_i, c_i denote the sides of n triangles $A_iB_iC_i$ ($i = 1, 2, \ldots\ldots., n$). Then the three numbers

$$a = \Sigma w_i a_i \qquad b = \Sigma w_i b_i \qquad c = \Sigma w_i c_i$$

where $\Sigma w_i = 1$, $w_i > 0$, are possible sides for a triangle ABC. Then,

$$K(ABC)^2 = \Sigma w_i s_i \,.\, \Sigma w_i (s_i - a_i) \,.\, \Sigma w_i (s_i - b_i) \,.\, \Sigma w_i (s_i - c_i)$$

and $s = \Sigma w_i s_i$. Applying Cauchy's inequality twice yields

$$\sqrt{K(ABC)} \geq \sum w_i \sqrt{K(A_i B_i C_i)}$$

with equality if and only if the n triangles are directly similar. Further inequalities can be obtained by averaging over the angles of the triangle [7].

8. Weierstrass Product Inequality.

In a recent problem-solving paper by Schoenfeld [12], there is a discussion about the following problem:

Let a, b, c, and d be given numbers in [0, 1]. Prove that

$$(1 - a)(1 - b)(1 - c)(1 - d) \geq 1 - a - b - c - d.$$

(We have changed the problem insignificantly by taking a closed interval rather than an open one.) Schoenfeld notes: "Virtually all of the Mathematicians I've watched solving it begin by proving the inequality $(1 - a)(1 - b) \geq 1 - a - b$. Then they multiply this inequality, in turn, by $(1 - c)$ and $(1 - d)$ to prove the three - and four-variable versions of it". Incidentally, this corresponds to Polya's heuristic, *specialize*. The proof for the n-variable version follows inductively. In a competition or on an exam, any method which works fairly quickly is just fine. However, if one wishes to get to the heart of the inequality, inductive proofs are usually not the way to go. This may be one of the subconscious reasons that many students shun such proofs.

An important property of the above inequality is that it is linear in each of the variables. Therefore, the inequality holds since it holds when each variable is at an endpoint of the interval [0, 1]. Not only is this proof more satisfactory than the inductive one, it leads to the following generalizations quite easily.

8.1 The right hand side of the n-variable inequality

$$(1 - x_1)(1 - x_2) \ldots (1 - x_n) \geq 1 - x_1 - x_2 - \ldots - x_n$$

(where $0 \leq x_i \leq 1$) is just the constant term plus the linear part of the product of the left hand side. So why stop with the linear terms! If we define $T_1, T_2, \ldots, T_n$ by

$$(t + x_1)(t + x_2) \ldots (t + x_n) \equiv t^n + T_1 t^{n-1} + T_2 t^{n-2} + \ldots T_n$$

then for $0 \leq x_i \leq 1$,

$$1 - T_1 + T_2 - \dots T_{2r} \geq \prod_{i=1}^{n} (1 - x_i) \geq 1 - T_1 + T_2 + \dots - T_{2s+1}$$

where $r = 0, 1 \dots [\frac{n}{2}]$, $s = 0, 1 \dots, \left[\frac{(n-1)}{2}\right]$ This last result is obviously valid if all the x_i are zero. If m (> 0) of the x_i are 1 and the rest are zero, then all we need show is that:

$$\genfrac{[}{]}{0pt}{}{m}{0} - \genfrac{[}{]}{0pt}{}{m}{1} + \dots + \genfrac{[}{]}{0pt}{}{m}{2r} \geq 0 \geq \genfrac{[}{]}{0pt}{}{m}{0} - \genfrac{[}{]}{0pt}{}{m}{1} + \dots - \genfrac{[}{]}{0pt}{}{m}{2s+1},$$

which follows by the unimodal and symmetric character of the binomial coefficients.

9. An Inequality.

If $1 \geq a, b, c, \geq 0$, then

$$\frac{a}{b+c+1} + \frac{b}{c+a+1} + \frac{c}{a+b+1} + (1-a)(1-b)(1-c) \leq 1$$

This inequality is a special case of a more general one due to Andre Giroux and first established by him using sophisticated methods. Subsequently he, and independently A. Meir and I, obtained simpler proofs. I then set the proof of the special case above as a problem in the 1980 U.S.A. Mathematical Olympiad. Most of the proofs given in the Olympiad were of the "direct brutal assault" kind not leading easily to extension. A simple proof follows by first noting that the function on the left hand side of the inequality is convex in each of the variables a, b, c. Thus the function takes on its maximum value at the extreme value 0 or 1 for each of the variables, i.e., at some vertex (a,b,c) of the cube whose coordinates are 0's and 1's. Since the value of the function is 1 at each vertex of the cube, we are done.

9.1 In a similar way, we can establish the more general inequality

$$\sum_{i=1}^{n} \frac{x_i^u}{1 + s - x_i} + \prod_{i=1}^{n} (1 - x_i)^v \leq 1$$

where $0 \leq x_i \leq 1$; $u,v \geq 1$; and $x_1 + x_2 + \dots + x_n = s$.
Giroux's inequality, corresponding to $u = v = 1$, later reappeared as a proposed problem [13].

10. Summations

Determine the sum $\frac{1}{2!} + \frac{2}{3!} + \ldots + \frac{n}{(n+1)!}$.

Schoenfeld [12] notes that most students will begin by doing the addition and placing all the terms over a common denominator, whereas the typical expert will calculate a few cases from which the inductive pattern is clear and easy. If all you want is the answer, this is okay. If you want more, then inductive proofs are not good enough as was indicated above in 8. I would also want the student to learn something about summation of series in general. Summation is analogous to definite integration, where by virtue of the fundamental theorem of calculus, one first tries to obtain the anti-derivative of the integrand. This search is simplified by having a table of integrals which is obtained in a reverse fashion. One merely takes a set of functions and differentiates them, which is an easy direct operation. Then one simply reverses the table to give the integral table. One can do the same thing with discrete sums. The fundamental theorem of finite summation is

$$\sum_{k=1}^{n} \{F(k) - F(k-1)\} = F(n) - F(0) .$$

So to find the above sum, "all we need to do" is express the summand as an "anti-derivative", i.e., as a difference $F(k) - F(k-1)$. Since the summand is

$$\frac{k}{(k+1)!} = \frac{k+1}{(k+1)!} - \frac{1}{(k+1)!} = \frac{1}{k!} - \frac{1}{(k+1)!} ,$$

the sum is immediately $1 - \frac{1}{(n+1)!}$

In general, however, finding the summation anti-derivative can be a difficult problem. To aid in this process, as in integration, we make a summation table in reverse fashion. This is done in books on finite differences, in particular see [14]. Here we start with (2) and make a table by considering various functions F. Finally, we reverse the order of the table. As examples, we have the following:

$$F(k) = (k+1)! \quad \text{yields} \quad \sum_{k=1}^{n} k.k! = (n+1)! - 1,$$

$$F(k) = \sin(ak + \frac{a}{z} - \frac{\pi}{2}) \quad \text{yields} \quad \sum_{k=1}^{n} \sin ak = \frac{\sin \frac{an}{2}}{\sin \frac{a}{2}} \, . \, \sin \frac{a(n+1)}{2},$$

$$F(k) = \sin(ak + \tfrac{a}{2}) \quad \text{yields} \quad \sum_{k=1}^{n} \cos ak = \frac{\sin \frac{an}{2}}{\sin \frac{a}{2}} \cdot \cos \frac{a(n+1)}{2},$$

$$F(k) = k^{m+1} \quad \text{yields}$$

$$\sum_{k=1}^{n} k^m = \frac{1}{m+1}\left\{ n^{m+1} + \begin{bmatrix} m+1 \\ 2 \end{bmatrix} \sum_{k=1}^{n} k^{m-1} - \begin{bmatrix} m+1 \\ 3 \end{bmatrix} \sum_{k=1}^{n} k^{m-2} \quad \ldots \right\}$$

Continuing the same way, we can obtain the sums

$$\sum_{n=1}^{\infty} \tan^{-1}\left[\frac{x}{n^2 + n + x^2}\right] = \tan^{-1} x,$$

$$\sum_{n=1}^{\infty} \frac{n}{2n^4 + 1} = \frac{1}{4}$$

Derivations of the latter two sums and related ones are given in [15].

It is to be noted that there are times when the inductive method does lead to an essential part of a given result. For example, consider the geometric theorem of T. Hayashi:

> If a convex polygon inscribed in a circle can be divided into triangles from one of its vertices, then the sum of the radii of the circles in these triangles is the same, whichever vertex is chosen.

An inductive proof leads one to consider the case of a quadrilateral first. In this case, it is the essential and hard part of the proof. The rest is easy.

For our final illustration, we consider an applied problem.

11. Flying in a Wind Field

It is a known elementary result that if an airplane flies an up and back straight course with a constant speed relative to a constant wind along the course, then the time of flight is greater than if there were no wind. One loses more time on the upwind part of the course then one gains on the downwind part, that is,

$$t = \frac{d}{v + w} + \frac{d}{v - w} = \frac{2vd}{v^2 - w^2} > \frac{d}{v} + \frac{d}{w}$$

where v = constant speed of the airplane, w = constant wind speed, $2d$ = total length of the course, t = time of flight.

11.1 To extend the previous result, let the airplane fly any path and then back along it in an arbitrary continuous stationary wind field. Here the wind speed and direction can vary in space but not in time. By considering the flight back and forth over any small element of the path and by using the above result, slightly modified, we obtain the same time inequality.

11.2 If we examine the expression for t above, we see that it is an increasing function of the wind speed, as is to be expected intuitively. To extend this result to 11.1, we replace the wind field $\vec{W}$ there by $k\vec{W}$ where k is a positive constant. We now wish to show that the time of flight in 11.1 is an increasing function of k.

First we obtain an expression for the total time of flight. Also, for comparison with a subsequent extension, we will assume that the path is a closed one. This is not necessary for the result here.

Let the arc length s denote the position of the plane on its path and let $w(s)$, $\theta(s)$ denote respectively the speed and the direction of the wind field $\vec{W}$ with respect to the tangent line to the path at position s. We will take the plane's speed as 1 and assume that $1 > kw$, otherwise the plane could not make the traverse. By resolving $k\vec{W}$ into components along and normal to the tangent line of the plane's path, the airplane's ground speed is

$$\sqrt{1 - k^2w^2\sin^2\theta} + kw\cos\theta$$

and then the time of flight is given by

$$T(k) = \oint \frac{ds}{\sqrt{1 - k^2w^2\sin^2\theta} + kw\cos\theta} + \oint \frac{ds}{\sqrt{1 - k^2w^2\sin^2\theta} - kw\cos\theta}$$

By the A.M. - G.M. inequality, the sum of the integrands is greater than or equal to $2(1 - k^2w^2)^{-1/2} \geq 2$ which shows that $T(k) \geq T(0)$ with equality if and only if $k\vec{W} = \vec{U}$. Then one can show that $T'(0) = 0$ and $T''(k) > 0$ since the integrand will consist solely of positive terms. Thus $T(k)$ is increasing in k.

11.3 The following nice extension was given by T.H. Matthews as a proposed problem [16]:

> *If an aircraft travels at a constant airspeed, and traverses (with respect to the ground) a closed curve in a horizontal plane, the time taken is always less when there is no wind, than when there is any constant wind.*

The solution given here leads immediately to another extension. If we let $\vec{W}$ be the wind velocity and $\vec{V}$ the actual plane velocity (which is tangential to the flight path), then $|\vec{V} - \vec{W}|$ is the constant speed of the airplane (without wind) and will be taken as unity for convenience.

We now have to show that

$$\oint \frac{ds}{|\vec{V}|} \geq \oint \frac{ds}{1}$$

By the Schwarz-Buniakowski inequality,

$$\oint |\vec{V}| ds) . \oint \frac{ds}{|\vec{V}|}) \geq (\oint ds)^2$$

Since

$$\oint |\vec{V}|\, ds = \oint \vec{V} \cdot d\vec{R} = \oint (\vec{V} - \vec{W}) \cdot d\vec{R} + \oint \vec{W} \cdot d\vec{R}$$

and

$$\oint \vec{W} \cdot d\vec{R} = \vec{W} \cdot \oint d\vec{R} = 0$$

we get

$$\oint |\vec{V}|\, ds \le \oint |\vec{V} - \vec{W}| \cdot |d\vec{R}| = \oint ds$$

(3) now follows from (4) and (5).

It is to be noted that in the above proof the closed horizontal curve can be replaced by a space curve and that the wind field $\vec{W}$ need not be constant; it can be irrotational since we still have

$$\oint \vec{W} . d\vec{R} = 0$$

11.4 The latter result can be extended as in 11.2. Let the wind field be $k\vec{W}$ where $\vec{W}$ is irrotational; then the time of flight is given by

$$T(k) = \oint \frac{ds}{\sqrt{1 - k^2 w^2 \sin^2\theta} + kw\cos\theta}$$

Then since $T'(0) = 0$ and $T''(k) > 0$, $T(k)$ is increasing in k.

For other related airplane problems, see [17, 18, 19, 20].

To become proficient in problem creating, as in any other non-trivial activity, one must have lots of practice. At the U.S.A. Mathematical Olympiad training sessions, I required the students not only to solve challenging problems but also to submit original reasonable proposed problems. I did this since I completely agree with Polya who points out in his books that these activities go hand in hand. Since most of these students had had little or no practice in problem creating, their first efforts were usually poor. However, with continual practice their submissions improved markedly. Here are two examples (for other ones, see [21]):

> A quick proof that the rationality of p, q and $\sqrt{p}+\sqrt{q}$ implies the rationality of $\sqrt{p}$ is furnished by the identity
>
> $$2\sqrt{p} = \frac{(\sqrt{p}+\sqrt{q})^2 + p - q}{\sqrt{p}+\sqrt{q}}$$
>
> Prove, in a similar fashion, that if p, q, r and $\sqrt{p}+\sqrt{q}+\sqrt{r}$ are rational, then so is $\sqrt{p}$.

(By Gregg Patruno, now a graduate student at Columbia University. He has recently extended the result to n rational numbers).

> Three disjoint spheres whose centers are not collinear are such that there exists eight planes each tangent to all three spheres. The points of tangency of each of these planes are vertices of a triangle. Prove that the circumcenters of these eight triangles are collinear.

(By Noam Elkies, now a graduate student at Harvard University).

Finally, to conclude this paper, I adapt a quotation of G. Polya.

> *"Proposing problems is a practical art like swimming or skiing, or playing the piano: you can learn it only by imitation and practice. This paper cannot offer you a magic key that opens all the doors and proposes all the problems, but it offers you good examples for imitation and many opportunities for practice; if you wish to learn swimming you have to go into the water, and if you wish to become a problem proposer you have to propose problems".*

References

1. *Crux Mathematicorum* 5 (1979) 220-227.

2. H. Rademacher and O. Toeplitz, *The Enjoyment of Mathematics,* Princeton University Press, Princeton, 1957, p. 139.

3. M.S. Klamkin, Vector proofs in solid geometry, *Amer. Math. Monthly,* 77 (1970) 1051-1065.

4. O. Bottema et al, *Geometric Inequalities*, Noordhoff, Groningen, 1969, p. 128.

5. M.S. Klamkin, On equilateral and equiangular polygons, *Crux Mathematicorum* 7 (1981) 2-5.

6. M.S. Klamkin, Intersecting curves on a surface, *Amer. Math. Monthly* 75 (1968) 876-877.

7. Problem 82-15, *SIAM Review* 25 (1983) 406-407.

8. Problem 84-9*, *Math. Intelligencer* 6 (1984) 39.

9. *Crux Mathematicorum* 6 (1980) 246-247.

10. M.S. Klamkin, Notes on inequalities involving triangles and tetrahedrons, *Univ. Beograd Publ. Electrotehn. Fak. Ser. Mat. Fiz.* No. 330 (1970) 1-15.

11. *Crux Mathematicorum* 10 (1984) 46-48.

12. A.H. Schoenfeld, Teaching problem-solving skills, *Amer. Math. Monthly* 87 (1980) 794-805.

13. Aufgabe 873, *Elem. der Math.* 37 (1982) 27.

14. C. Jordan, *Calculus of Finite Differences*, Chelsea, New York, 1947, p. 104.

15. M.L. Glasser and M.S. Klamkin, On some inverse tangent summations, *Fibonacci Quarterly* 14 (1976) 385 - 388.

16. Problem 4132, *Amer. Math. Monthly* 52 (1962) 584-585.

17. M.S. Klamkin and D.J. Newman, Flying in a wind field I, *Amer. Math. Monthly* 76 (1969) 16-23.

18. M.S. Klamkin and D. J. Newman, Flying in a windfield II, *Amer. Math. Monthly* 76 (1969) 1013 - 1019.

19. M.S. Klamkin, On extreme length flight paths, *SIAM Review* 18 (1976), 486 - 488.

20. M.S. Klamkin, On Chaplygin's problem, *SIAM J. Math. Anal.* 8 (1977), 288 - 289.

21. *Crux Mathematicorum* 6 (1980) 210; 8 (1982) 99 - 100, 300 - 301; 10 (1984) 74.

* * *

Crux Mathematicorum, published by the Canadian Mathematical Society, is a problem solving journal at the senior secondary and university undergraduate levels for those who practice or teach mathematics. Its purpose is primarily educational, but it serves also those who read it for professional, cultural or recreational reasons.
It is published monthly (except July and August). The yearly subscription rate for ten issues is $C22.50 for members of the Canadian Mathematical Society and $C25 for non-members.Cheques and money orders, payable to Crux Mathematicorum, should be sent to:-

The Managing Editor,
Dr Kenneth Williams,
Canadian Mathematical Society,
577 King Edward Avenue,
Ottawa, Ontario,
CANADA, KIN 6N5.

* * *