Invariants and monovariants

Juni Drakengren, Ludvig Granath, Nils Kalmnäs Drakenfors,
Olle Rehnquist, Erik Wettergren

Hvitfeldtska Gymnasiet

May 2022

Contents

1 Introduction

One of the fundamental principles of mathematical problem solving is the detection of patterns. Arguably, the most simple pattern conceivable is that of constancy. An unchanging variable may be indicative of some underlying framework. Understanding such fundamental frameworks can often transform a seemingly difficult problem into something which is essentially trivial. Unchanging properties of a problem are denoted *invariants.* This essay will survey the use of invariants in a wide range of mathematical contexts: from combinatorics, to game theory, inequalities and geometry. First of all, however, we must formalise what is meant by "invariant".

1.1 Invariants

An invariant is any property of a mathematical object which remains unchanged under a certain transformation. This definition may seem vague and unspecific. But as we will come to see, its generality is necessitated by the diverse contexts in which invariants are useful. However, in interest of providing better grounds for understanding the idea, let us first examine a concrete example:

Example 1.1. Consider a chessboard with its top-left and bottom-right tiles removed. Imagine now that we place domino-pieces on the chessboard. The dominoes must cover exactly two squares each, either horizontally or vertically, and they may not overlap. Given these prerequisites, it it possible to tile the entire chessboard with dominoes?

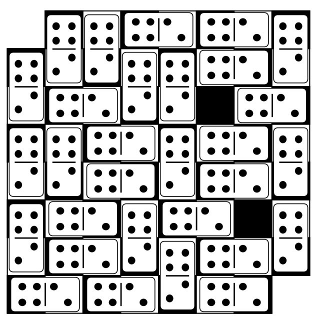


Figure 1: A (failed) tiling of the given chessboard using domino-pieces.

Solution. Let us think of the problem is terms of invariants. Does the chessboard have any property which remains unchanged when we remove two adjacent squares? We notice that adjacent squares are always of different colour. This means that even as we place dominoes, the relative difference between the number of black and white squares will always be the same. It is an invariant. Now, both the corner pieces which we removed were white. This means that we start with two more black squares than white. Consequently, due to the invariant, there will always be two more black squares than white. If we reach a tiling where only two squares remain, both of these must therefore be black (as in Figure 1 above). But two black tiles are never present next to each other, so we can never place a domino to cover the two final squares. The tiling is impossible. □

Example 1.2. A zombie plague is infecting a town. Each night, the townspeople will meet randomly in groups of three. If at least one of the members of the group is a zombie, all others will become zombies at the end of the night. Given that the infection starts in one random individual, and that the town has a population of 1000, can there be exactly 100 affected individuals after 10 nights?

Solution. By thinking in terms of invariants, the solution becomes quite simple. We notice that since the groups are always composed of three individuals, each group with at least one zombie will have exactly three zombies at the end of each night. Thus, the number of zombies at the end of each night is invariably divisible by 3. Since 3 does not divide 100, there cannot ever be 100 affected individuals. □

1.2 Monovariants

Just as constancy is a fundamental pattern in problem solving, so is change. In particular, we may be interested in mathematical properties which always change in the same way. This is where the idea of monovariants comes in. A monovariant is a property which changes monotonically – that is, in a predictable way. The interpretation of ”predictable” may vary depending on the context. One common take is variables whose value only increases or decreases.

Example 1.3. Given an even amount of points in the plane, where no three points lie on the same line, is it possible to divide them into pairs connected by non-intersecting line segments?

Solution. Begin by randomly pairing up all points and connecting these pairs. Now, pick any two such pairs of points $(A_1, B_1), (A_2, B_2)$ whose line segments $\overline{\mathrm{A_1B_1}}$ and $\overline{\mathrm{A_2B_2}}$ intersect. Given this, it must be that $\overline{\mathrm{A_1B_2}}$ and $\overline{\mathrm{B_1A_2}}$ do not intersect. Erase the original segments and replace them with the non-intersecting ones. According to the triangle inequality, the total length of the line segments has strictly decreased. Now, repeat this process. For each step, the total length of all of the line segments decreases. It is a monovariant. Because of this, and the fact that there is a finite amount of ways to connect all of the points to each other, the process cannot continue indefinitely. When the process terminates, there will no longer be any intersections. □

Having covered the essentials of invariants and monovariants, we now survey their usage in more specific contexts.

2 Combinatorics

Invariants are most prominently used in combinatorics, the study of counting. The subject of this count are most often and either finite arrangements, combinations or permutations. Intertwined in this very foundation of combinatorics is one invariant of high importance: For a count of the objects in a finite set, regardless of the order in which the objects are counted, the end result will be the same. Even outside of this fundamental invariant, many combinatorial problems contains other invariants or monovariants of utmost usefulness when searching for a solution.

Example 2.1. In a tournament of n teams, where each team consists of t players, every possible pair of two players from the same team plays exactly one game against every other possible pair of players from all other teams. Find the number of games played in the tournament.

Solution. Every game will consist of two pairs of players from two different teams. Notice that since all teams contains the same number of players, and therefore the same number of possible pairs, the number of games played between pairs from any two teams is an invariant.

To find the number of games played in the tournament we need to find the number of combinations of two teams, and the number of games played by pairs from two fix teams. The former is calculated by $\binom{n}{2}$. Calculating the latter is similarly easy. The number of possible pairs is counted just like the number of pairs of teams, by $\binom{t}{2}$, which gives that the number of games played between two fix teams is $\binom{t}{2}^2$.

We have calculated the number of combinations of two teams to be $\binom{n}{2}$, and the amount of games played between two teams to be $\binom{t}{2}^2$. The number of games played in the tournament is the product of these two numbers: $\binom{n}{2}\binom{t}{2}^2$ □

In this example an invariant was used to simplify a basic, but decently complex, problem into two simple combinatorial calculations. Although it was the case here that the invariant was pretty easy to find, that is far from always the case. An example of where that is quite tricky is problem 5 in the 2021/2022 SMT finals, where a partition is found so that all sets have a common trait.

Example 2.2. Let n be a positive integer congruent with 1 modulo 4. Xantippa has a bag with $n+1$ balls numbered from 0 to n. She pulls a random ball from the bag and reads its number: k. She keeps the ball and pulls k additional random balls from the bag. Finally she sums up all the numbers on the $k+1$ balls she has pulled. What is the probability that the sum is odd?

Solution. There are a few different paths towards finding the probability, though all goes through an invariant. The key is noticing the similarities between pulling any ball, numbered s, and the corresponding ball numbered $n-s$. If the ball numbered $n-s$ is pulled, $n-s+1$ balls will be pulled before the summation, leaving $(n+1)-(n-s+1)=s$ balls in the bag. So for each ball s which can be pulled, there is one other ball $n-s$ which when pulled leaves exactly s balls in the bag. However, for the ball $n-s$, the ball which when pulled leaves $n-s$ balls in the bag turns out to be $n-(n-s)=s$. This means that all balls in the bag forms this kind of a pair with one other ball. (No ball would form a pair with itself as $n-s$ never equals s due to n being odd.)

As $(n-s)+s=n$, which is odd, one ball, v, in each pair is even and the other, d, is odd. After the initial ball has been pulled, the numbers on the balls is no longer relevant outside of its parity, as all numbers of one parity affects the parity of the sum in the same way. Because the bag initially contained an equal amount of odd and even balls, when v is pulled there remains one more odd than even ball in the bag, and the opposite will be the case for d. Due to the reason for why the two balls were paired, if v is pulled there will remain d balls in the bag, and of course if d is pulled d more balls will be drawn. Combined these conclusions means that the probability for every specific combination D of *odd and even* balls after pulling the ball d, must equal the probability of leaving in the bag the reversed combination V of *even and odd* balls, after pulling v. That is to say, that V has the same number of odd balls as D has even balls, and V has the same number of even balls as V has odd balls. Due to both combinations containing d, an odd number, of balls, this implies that the combinations have different parity . Additionally since the sum of all the balls are odd, the sum of the pulled balls has a different parity than the sum of the balls remaining in the bag. This gives that all balls pulled from the bag when V was left in the bag (including the initially pulled v), say V_p, has the opposite parity of V, and therefore the same parity as D. But D is only the balls pulled after d was initially pulled, so the sum of all the balls pulled with d, say D_p, actually has the opposite parity of D and V_p.

This means that for every combination of balls pulled from the bag after an initial ball s, there is one unique corresponding combination pulled after the other ball in its pair, $n-s$, that is equally likely and of the opposite parity. But if there for every possible combination of pulled balls after s is one corresponding equally likely combination pulled after $n-s$ which gives the opposite parity, then that must mean that the probability of the sum being odd after initially pulling s equals the probability of the sum being even after initially pulling $n-s$, and as it is equally likely to pull both balls, with the information that a ball from a pair $(s, n-s)$ was pulled, the probability of the sum being odd is exactly 50 percent. As every ball is in a pair, that information is always held and the probability of the sum being odd is exactly 50 percent. ☐

2.1 Colouring Proofs

Another common method for solving combinatorial problems is the colouring proof. This is a type of proof which uses some kind of colouring on a board or tiles, like in example 1.1. The goal of the colouring is essentially always to find an invariant, though its nature can wary significantly. We will not touch this method in depth, but instead recommend the paper *The method of colouring in combinatorics*[1], written by Rosendalsgymnasiet in this same event two years prior.

Example 2.3. A board is made up of 44 equally sized squares and has the shape of a rectangle of dimensions 4:11. Is it possible to completely fill the board with L-shaped pieces that each cover four tiles? Pieces may not overlap, but may be flipped, turned and moved in any way.

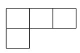
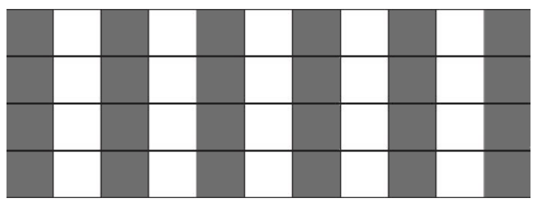


Illustration of the L-shaped piece.

Solution. We colour every second column grey:



We can now see that every piece will either cover three grey square and one white square, or three white square and one grey square, regardless of its placement. To fill every square, 11 pieces are needed, due to the sizes of the board and the pieces. However, as each piece fills an odd number of tiles of each colour, 11 tiles will also fill an odd number of tiles of both colours, but the board contains an even number of tiles of both colours and it is therefore impossible to completely fill the board with the pieces. ☐

2.2 Problems

Problem 2.1 During an event 37 players plays one on one games of chess. There is no limit to the amount of games a player can play, nor to the amount of times two players can play against one another. Show that at the end of the event, the number of players who played an even number of games is odd.

Problem 2.2 Initially, 9 of the 100 squares in a 10×10 grid are infected. During each unit time interval, each square which has 2 or more infected neighbours (a neighbour being a square which shares an edge) also becomes infected. Determine whether it is possible that all 100 squares will eventually become infected.

[1] See chapter 6: Further reading

Problem 2.3 A board made up of equally sized squares is in both dimensions infinite but limited in one direction. Each square has unique coordinates on the form $(a,\ b)$, where a, b are integers ≥ 1. Three stem cells exists on the squares (1, 1), (2, 1) and (1, 2) respectively. A stem cell on any square $(a,\ b)$ can at any time divide into two stem cells, that appears on squares $(a{+}1,\ b)$ and $(a,\ b{+}1)$, as long as neither of those two squares are occupied by a stem cell. Is it possible after a finite number of divisions that neither of the squares (1, 1), (2, 1) and (1, 2) are occupied by a stem cell?

3 Game theory

Game theory as a field is mainly concerned with two kinds of problems: determining whether or not a given game has an optimal strategy, and finding an optimal strategy (note that this is also a way to determine the former). Invariants and monovariants prove highly useful in both endeavours.

3.1 Invariants and monovariant in game theory

Since game theory concerns itself with theoretical rational actors, invariants are useful for finding limits to the results of their actions; since invariants are, per definition, properties that remain unchanged, they allow us to study properties of the game without knowing its optimal strategy. Likewise, monovariants show which parameters of the game change predictably. Consider the following examples:

Example 3.1. Two players take turns breaking up an $m \times n$ chocolate bar. On a given turn, a player picks a rectangular piece of chocolate and breaks it into pieces along the subdivisions between its squares. The player who makes the last break wins. Does one of the players have a winning strategy?

solution. Let A be the number of pieces of chocolate. Initially, we find that $A = 1$. In the end, regardless of how the game is played, we find that $A = n \times m$. For every new move, the number of pieces increases by exactly 1. This, in turn, means that the game results in victory for the first player solely if $n \times m - 1$ is odd, and in victory for the second player if it is even. There is therefore no strategy that guarantees any player victory.

□

Note that in the example above, it is shown that the length of the game is in invariant by combining a monovariant (the number of pieces of chocolate) and another invariant (the game always terminates in the same manner). This method can also be applied to more complicated problems, like the following:

Example 3.2 (Nordic Mathematical Contest, 2014, problem 4)**.** A game is played on an $n \times n$ chessboard. At the beginning there are 99 stones on each square. Two players A and B take turns, where in each turn the player chooses either a row or a column and removes one stone from each square in the chosen row or column. They are only allowed to choose a row or column, if it has at least one stone on each square. The first player who can't make a move looses the game. For which n does A have a winning strategy?

Solution. Assume that the game is over, but the board is not empty. Then there must be some square (a, b) such that there are still stones on it, but at least one empty square in the same row (a, c) and one empty square in the same column, (d, b). Let R_i denote the number of times a stone has been removed from every square in row i ad C_j the number of times a stone has been removed from every square in column j. We see that $R_a + C_b < 99$ (since there are still stones left in square (a, c)) and that $R_a + C_c = 99$, $R_d + C_b = 99$. However, this leads to a contradiction, since this

means that $R_d + C_c > 99$, in turn meaning that more than 99 stones must have been removed from square (d, b), which is impossible.
The game must therefore always end the same way, regardless of how it is played, in turn meaning that the length of the game is invariant (since the same number of stones are removed every time a player makes move, leading to a game of length $99n^2/99 = n^2$). This in turn means that player A wins if, and only if n is odd. $\square$

As shown above, the use of invariants and monovariants greatly simplifies our study of games, since they ensure that the particulars of optimal play can be dealt with later, or completely ignored altogether.

3.2 Problems

Problem 3.1 Let $n \geq 2$ be a positive integer. The game starts with $2n$ piles of stones. Each pile has exactly one stone. A and B alternately merge two piles of their own choice, until there are only two piles left. A wins if those piles have even numbers of stones each, and B if the number of stones in each pile is odd. A starts the game. Who has a winning strategy, and what is it?

Problem 3.2 (*Baltic Way*, 2021) Let $n > 2$ be an integer. Anna, Edda and Magni play a game on a hexagonal board tiled with regular hexagons, with n tiles on each side. The game begins with a stone on a tile in one corner of the board.

Edda and Magni are on the same team, playing against Anna, and they win if the stone is on the central tile at the end of any player's turn. Anna, Edda and Magni take turns moving the stone: Anna begins, then Edda, then Magni, then Anna, and so on.

The rules for each player's turn are:
• Anna has to move the stone to an adjacent tile, in any direction.
• Edda has to move the stone straight by two tiles in any of the 6 possible directions.
• Magni has a choice of passing his turn, or moving the stone straight by three tiles in any of the 6 possible directions.

Find all n for which Edda and Magni have a winning strategy.

4 Inequalities

Invariants frequently occur in the study of inequalities and within this section we will develop some intuition and general techniques for finding them.

4.1 Scaling variables

The most simple example of invariants in inequalities is when an inequality is preserved under scaling of the variables by a positive real number.

Example 4.1. Prove that if $0 < p < 1$ and x, y are positive real numbers, then $(x+y)^p < x^p + y^p$.

Solution. Notice that if we let x map to λx and y map to λy where λ is a positive real number, we obtain the inequality

$$\lambda^p (x+y)^p < \lambda^p (x^p + y^p) \iff (x+y)^p < x^p + y^p.$$

Now we can assume that $x + y = 1$ since scaling both variables by a positive number gives an equivalent inequality and so we ought to prove that $1 < x^p + y^p$. But because $x, y < 1$ (since their sum is 1) we have that $x^p > x$ and $y^p > y$, therefore $x^p + y^p > x + y = 1$ which was what we wanted to prove. □

We saw that it might be sufficient to prove the inequality when we have imposed a restriction on the variables if the inequalities are equivalent. Next we will demonstrate an interesting use of invariants for proving the famous Cauchy-Schwarz Inequality.

Example 4.2 (Cauchy-Schwarz Inequality)**.** For all sequences of real numbers a_i and b_i it holds that

$$(\sum_{i=1}^{n} a_i^2)(\sum_{i=1}^{n} b_i^2) \geq (\sum_{i=1}^{n} a_i b_i)^2.$$

Proof. We get an equivalent inequality by exchanging a_i with λa_i for a positive real number λ. Therefore we can assume without loss of generality that $\sum_{i=1}^{n} a_i^2 = 1$ and in the same way that $\sum_{i=1}^{n} b_i^2 = 1$. The only special case where this is not valid is if $\sum_{i=1}^{n} a_i^2 = 0$ or $\sum_{i=1}^{n} b_i^2 = 0$, but then $a_i = 0$ or $b_i = 0$ for all i and we can check that Cauchy-Schwarz inequality holds in this case. Otherwise the inequality reduces to

$$1 \geq (\sum_{i=1}^{n} a_i b_i)^2 \iff 1 \geq |\sum_{i=1}^{n} a_i b_i|.$$

But we can here use the AM-GM inequality (refer to example 4.5).

$$1 = \sum_{i=1}^{n} \frac{a_i^2 + b_i^2}{2} \geq \sum_{i=1}^{n} \sqrt{a_i^2 b_i^2} = \sum_{i=1}^{n} |a_i b_i| \geq |\sum_{i=1}^{n} a_i b_i|.$$

The last step follows from the triangle inequality. □

4.2 Derivative

Another invariant that we can consider in some problems is the derivative of an expression. We will begin with a very simple example.

Example 4.3. Prove that $e^x > x + 1$ for $x > 0$.

Solution. Set $f(x) = e^x - x - 1$, then we want to show that $f(x) > 0$ for $x > 0$. But notice that the sign of the derivative $f'(x) = e^x - 1$ is invariantly positive for $x > 0$, so the function is strictly increasing (a monovariant). Since $f(0) = 0$ the claim follows. □

We will now consider a more ingenious application of invariants of the derivative which exemplifies well the diversity of the concept.

Example 4.4 (British Mathematical Olympiad 1995 Round 2 Problem 3)**.** Let a, b, c be real numbers satisfying $a < b < c, a + b + c = 6$ and $ab + bc + ca = 9$. Prove that $0 < a < 1 < b < 3 < c < 4$.

Solution. Since the requirements for a, b, c look very similar to Vieta's formulae we immediately deduce that a, b, c are zeroes of the polynomial function

$$p(x) = x^3 - (a + b + c)x^2 + (ab + bc + ca)x - abc = x^3 - 6x^2 + 9x - abc.$$

We see that only the constant term is dependent on a, b, c, hence we understand that the derivative of p is invariant of a, b, c. More precisely

$$p'(x) = 3x^2 - 12x + 9 = 3(x-3)(x-1).$$

Because p initially is increasing it changes sign at $x = 1$ and becomes negative and changes to positive at $x = 3$ again. The proof is intentionally left incomplete for the reader to solve the problem using this fact, which should be trivial. $\square$

With more refined mathematical analysis, many problems can be trivialised by invariants, though these are beyond the scope of this text.

4.3 Monovariants in inequalities

We have already seen some monovariants and the distinction between them and invariants is often unnecessary, but this subsection will treat a classical problem solving technique.

Example 4.5. (AM-GM) Prove that for all non-negative numbers $x_1, x_2, \dots, x_n$ it holds that

$$\frac{x_1 + \cdots + x_n}{n} \geq \sqrt[n]{x_1 \cdots x_n}.$$

Proof. Consider given numbers $x_1, \dots, x_n$ and suppose that not all of them are equal to the arithmetic mean a. Then for some i, j we have that $x_i < a < x_j$ (can you see why?). We now let

$$x_i' = a$$
$$x_j' = x_i + x_j - a$$

Notice how the sum $\frac{x_1+\cdots+x_i'+\cdots+x_j'+\cdots+x_n}{n}$ has not changed (since $x_i' + x_j' = x_i + x_j$) but

$$x_i'x_j' = a(x_i + x_j - a) = x_ix_j + (a - x_i)(x_j - a) > x_ix_j.$$

Therefore the right hand side of the inequality has increased strictly but not the left hand side - this is our invariant. Now this process of making the right hand side larger will terminate, as we have a finite number of variables and then all of them will be equal to the arithmetic mean. If $x_1 = \cdots = x_n$ we see that equality holds. $\square$

This very process is sometimes referred to as *smoothing*. Another strategy to keep in mind is to increase one side of an inequality while retaining the value of the other side, which we did above.

4.4 Problems

Problem 4.1 If a, b, c are positive real numbers, prove that

$$\frac{a}{(b+c)^2} + \frac{b}{(c+a)^2} + \frac{c}{(a+b)^2} \geq \frac{9}{4(a+b+c)}.$$

Problem 4.2 Prove that if $a, b > 0$ then

$$(\frac{a+1}{b+1})^{b+1} \geq (\frac{a}{b})^b.$$

Problem 4.3 Given real numbers $x_1, \ldots, x_n$, what is the minimum value of

$$|x - x_1| + \cdots + |x - x_n|?$$

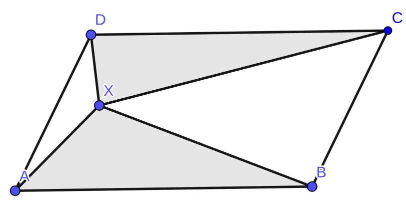
5 Geometry

In Geometry, common invariant-problems involve geometric transformations like for example isometries, including reflections and translation, operations which preserves many properties. For more on geometric transformations, see chapter 6: Further reading. In some problems, there might also be occurrences of area-invariants.

5.1 Invariant area

Example 5.1. Let $ABCD$ be a parallelogram, and X a point inside the parallelogram. Find the minimum sum of the areas of $\triangle ABX$ and $\triangle CDX$ expressed in terms of a and h, where $|AB| = a$ and h is the distance between sides CD and AB.

Solution. To begin with, we consider the case where X coincides with A. Calculating the sum of the areas of $\triangle ABX$ and $\triangle CDX$ (coloured grey in the picture below) becomes easy as it is half the area of the parallelogram, that is $\frac{a \cdot h}{2}$. If we can prove that the area is invariant under the operation of moving the point X inside the parallelogram, we have found the right answer.

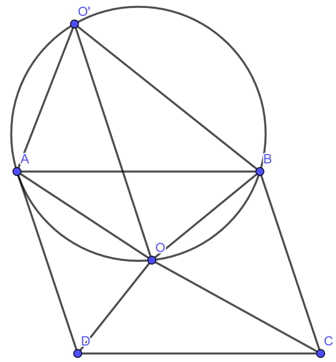


Moving X can be seen as a combination of two operations; moving the point parallel to side AB, and moving X orthogonally to AB. In the first case, we can see that both the altitudes and the bases of the grey triangles are constant, which means the areas of the grey triangles are constant. In the second case, both the sum of the length of the altitudes of the grey triangles, and the bases remain unchanged. Denote the altitudes of $\triangle ABX$ and $\triangle CDX$ with h_1 and h_2 respectively. The sum of the areas can now be calculated by the following formula: $\frac{a \cdot h_1}{2} + \frac{a \cdot h_2}{2}$. But as $h_1 + h_2 = h$ this is equivalent to $\frac{a \cdot h}{2}$, which hence is the minimal area. $\square$

5.2 Translation

Example 5.2. (*Canada* 1997) The point O is situated inside the parallelogram $ABCD$ such that $\angle AOB + \angle COD = 180°$. Prove that $\angle OBC = \angle ODC$.

Solution. Since $ABCD$ is a parallelogram, we can translate $\triangle DOC$ to triangle $\triangle AO'B$ as in the picture below.



As both distances and parallelity are preserved under translation, it follows that $|OO'| = |DA| = |BC|$, and that $OCBO'$ is a parallelogram. Using the parallel postulate, we receive that $\angle OBC = \angle O'OB$.

Also angles are preserved under translation, which means that $\angle O'AB = \angle ODC$ and that $\angle AO'B + \angle COD = 180°$. Consequently, quadrilateral $AOBO'$ is cyclic. By the inscribed angle theorem, we know that $\angle O'AB = \angle O'OB$. Using these facts we can now conclude that $\angle OBC = \angle O'OB = \angle O'AB = \angle ODC$. □

5.3 Problems

Problem 5.1 (*SMT* 2019/2020) The segment AB is the diameter in a circle with radius R. Point P lies on the diameter. The quadrilateral $PBCD$ is a rectangle such that line CD is tangent to the circle. Segment DP intersects the circle in Q. Given that $|QB| = 6cm$,
a) find the smallest R_0 such that the construction is possible for every $R \geq R_0$
b) show that the area of rectangle $PBCD$ for $R \geq R_0$ does not depend on R and calculate this area. [2]

Problem 5.2 (*SMT 2018/2019*) Let AB be a chord in a circle with centre O. Line l passes through O and intersects the chord AB in point P. Let C be the reflection of B in line l. Show that points A, C, O and P lies on a circle.

6 Further reading

As we have seen, invariants occur in a wide variety of settings and unfortunately we can not cover them all in this introductory essay. For further reading the reader is advised to read the following literature:

- *Problem-Solving Strategies* by Arthur Engel, and especially its first chapter.
- *Euclidean Geometry in Mathematical Olympiads* by Evan Chen.
- *The method of colouring in combinatorics* by V. Jansson, V. Laaksoharju, D. Mörtberg, P. Nilsson and F. Rönngren, as was already mentioned in section 4.

[2]This problem uses a monovariant.