Reduction to Absurdity

Li-Chung Wang

February 28, 2011

Abstract

First, if an argument that uses reduction to absurdity, we want to make the argument more direct or straightforward. Suppose $A \Rightarrow C$ is proved using reduction to absurdity. In order to make our argument more straightforward, we try to seek a middle statement B so that either $(A \Rightarrow B)$ or $(B \Rightarrow C)$ can be proved without using reduction to absurdity. In this sense, Galois' proof of the Abel–Ruffini Theorem is more straightforward than Abel's proof; the elementary proof of the prime number theorem is more straightforward than the transcendental proof; Jacobson's and Nelson's proofs of the Jordan normal form theorem is more straightforward than Herstein's. Second, we discuss why a direct proof is better than a proof using reduction to absurdity. There is no way for a proof using reduction to absurdity to change its feature of crudeness, indirectness, and limitedness; in contrast, a direct proof may reveal more amount of related information with less effort and continue to allow the improvement of precision for its conclusion whenever possible. Third, we define the reduced length and the total length of an argument using reduction to absurdity. Fourth, we discuss the embedded levels of reduction to absurdity. Fifth, we study the intrinsic nature of reduction to absurdity (How an artifical definition of rigidity necessitates the use of reduction to absurdity in the proof of Chasles' theorem. A proof using reduction to absurdity often provides only a local view and a shallow level of meaning; only through a straightforward proof may we have a global view, get to a deep level of meaning, and thus hit the heart of the matter. A physical proof is usually more direct than a geometric proof) and how we take advantage of the method (Choosing the method that can eliminate most impossibilities, investigating the cases that a contradicton surely occurs on an assigned path), and figure out some general methods of shortening the length of an argument using reduction to absurdity (if the original proof is essentially straightford, we should avoid giving it an outlook of a reduction to absurdity). The first method starts from a microscopic point of view. We should thoroughly explore the properties of our solution and fully take advantage of available resources. The more specialized cases we consider, the more resources we have. The more deeply we study the subject, the more confusions we may clarify and the more chances we have to find a shortcut for solutions. Our goal is to find a target for contradiction which has a simple, basic and close relationship with the theme of the theorem to be proved. A simple, basic, and close relationship may help us avoid an argument detour and identify the main factor that leads to the contradiction. We may replace a distant target with a nearby target by tracing the source of contradiction. Building a dense flow chart of theorems for a mathematical subject helps prove Theorem A using reduction to absurdity. This is because finding a theorem with 90% similarity to Theorem A becomes a simple task if we use the table of contents or the method of organization as our flow chart. The incompatibility of a false assumption and the limited differences between two theorems should help us quickly obtain a contradiction by comparison. The second method starts from a macroscopic point of view. We should use the Venn diagrams as our guide to understand the situation and design our depolyment. Then we should choose the weakest target for contradiction: Suppose $A \Rightarrow B$ and we choose B as the target for contradiction. In order to obtain a contradiction, we must prove the statement "*notB*". Thus, we must consider the complement of B, which is smaller than the complement of A in the Venn diagram. Having a smaller scope for consideration is the reason why choosing a weaker target for contradiction makes it easier to obtain a contradiction than choosing a stronger one. If we can try to proceed with our argument by both methods, they together may help us find the most direct argument. Sixth, we discuss how we shorten the total length for an argument using reduction to absurdity. Seventh, the concept of straightforwardness can be used to determined a proof's quality. Conclusion: The final goal of mathematicians is to prove any theorem it without using reduction to absurdity; in calculus of variation, some reduction-to-absurdity argument that seems unlikely to be made straghtforward can be executed with the method of Lebesgue integration. Even though a theorem involves the axiom of choice, we can still make its reduction-to-absurdity proof straightforward.

Keywords. Reductio ad Absurdum, reduction to absurdity, direct or straightforward, reduced lengths, total lengths, embedded levels, intrinsic nature, targets for contradiction, Venn diagrams, the Abel–Ruffini Theorem, the prime number theorem, the Jordan normal form theorem, a reduced pair of fundamental periods, nilpotent transformations, Mathieu's equation, a Flouquet solution, and the Sturm comparison theorem, Lebesgue integration, calculus of variation, Laurent series, meromorphic functions, generators of a one-sheet hyperboloid, Grönwall's inequality, the axiom of choice, Kolmogorov extension theorem, compact class, outer measure

1 How we make an argument using reduction to absurdity more straightforward

Suppose we prove the statement $A \Rightarrow C$ using reduction to absurdity. In order to make our argument more straightforward, we seek a middle statement B so that either ($A \Rightarrow B$) or ($B \Rightarrow C$) can be proved without using reduction to absurdity.

Example 1.1.

Ireland [22, p.145, l.6–l.17] is more straightforward than Ireland [22, p.144, l.15–l.27] because the former reveals such a middle statement: $p|N_f$.

Example 1.2.

Let
Statement A_1 = "$f = -\zeta'/\zeta$",
Statement A_2 = "$B(x) \sim 2x\, log\, x$" (Ellison–Ellison [12, p.93, l.-2]),
Statement A_3 = "g(s) does not have any multiple poles on $Re\, s = 1$ except for $s = 1$" (Elliso–Ellisonn [12, p.94, l.6]),
Statement A_4 = "(f has only simple poles on $Re\, s = 1$) and

$$Res(f;s_0) = \begin{cases} -1 & \text{if } s_0 \neq 1 \\ 1 & \text{if } s_0 = 1 \end{cases} \text{" (Ellison–Ellison [12, p.93, Theorem 3.9]),}$$

and
Statement A_5 = "f has no poles on $\{s|Re\, s = 1 \text{ and } s \neq 1\}$".

Ellison–Ellison [12, p.55, l.3] says that the Prime Number Theorem is a simple consequence of Ellison–Ellison [12, p.54, Theorem 2.10]. Ellison–Ellison [12, p.54, Theorem 2.10] can be derived from Ikehara's theorem (Ellison–Ellison [12, p.56, Theorem 2.11]) whose proof uses the fact $A_1 \Rightarrow A_5$ (Ellison–Ellison [12, p.91, l.-8–p.92, l.2]). We may use a transcendental method (Ellison–Ellison [12, p.41, l.-9–l.-1]) to prove $A_1 \Rightarrow A_5$. Thus, the only reduction to absurdity we use to prove the Prime Number Theorem lies in Ellison–Ellison [12, p.41, l.-9–l.-1]. In contrast, Erdös and Selberg insert three statements between A_1 and A_5. In

fact, they prove that A_1
$\Rightarrow A_2$ (Ellison–Ellison [12, p.96, Theorem 3.10])
$\Rightarrow A_3$ (Ellison–Ellison [12, p.94, l.7–l.14])
$\Rightarrow A_4$ (Ellison–Ellison [12, p.94, l.15–l.-2])
$\Rightarrow A_5$ (Ellison–Ellison [12, P.95, l.3–l.-1]).

$A_1 \Rightarrow A_2$ and $A_3 \Rightarrow A_4$ are proved without using reduction to absurdity. The reduction to absurdity used in proving $A_2 \Rightarrow A_3$ can be considered trivial. Thus, the elementary proof of the Prime Number Theorem is more straightforward than the transcendental proof.

2 Why a direct proof is better than a proof using reduction to absurdity

The drawbacks of a proof using reduction to absurdity: 1. It is not intuitive because of its indirectness; 2. It is difficult for its intrinsic crudeness to satisfy the requirement of rigor and precision in mathematics. In other words, there is no way for a proof using reduction to absurdity to change its feature of crudeness and indirectness. In contrast, a direct proof continues to allow the improvement of precision for its conclusion whenever possible. 3. The information that proofs using reduction to absurdity can reveal are limited, while a direct proof may reveal the most amount of related infomation with the least effort.

Example 2.1. (Linearly independent solutions of the Bessel equation)
When ν is not an integer, prove that $J_\nu(z)$ and $J_{-\nu}(z)$ are linearly independent solutions of the Bessel equation.

Direct Proof. In order to prevent the denominator of the right side of the equation given in Watson [42, p.40, l.3] to be 0, we assume that 2ν is not an integer [Watson [42, p.40, l.19]]. In this case, $J_\nu(z)$ and $J_{-\nu}(z)$ are solutions of the Bessel equation [Watson [42, p.40, l.-14–l.-13]]. When ν is half of an odd integer, we may use the argument given in Watson [42, §3.11] to prove that $J_\nu(z)$ and $J_{-\nu}(z)$ defined according to Watson [42, p.40, (8)] are still solutions of the Bessel equation. Then use the argument given in Watson [42, §3.12] to prove Watson [42, p.43, (2)]. □

Proof using reduction to absurdity. When 2ν is not an integer, $J_\nu(z)$ and $J_{-\nu}(z)$ are linearly independent solutions of the Bessel equation [Guo–Wang [17, p.348, l.8–l.10]]. However, if ν is not an integer, but 2ν is an integer, Guo–Wang [17] gives no proof. The argument using reduction to absurdity may apply to the case when ν is not an integer, but 2ν is an integer, but we must first use the argument given in Watson [42, §3.11] to prove that $J_\nu(z)$ and $J_{-\nu}(z)$ defined according to Watson [42, p.40, (8)] are still solutions of the Bessel equation when ν is half of an odd integer. However, Guo–Wang [17, §7.3] expresses $J_\nu(z)$ and $J_{-\nu}(z)$ as series involving cos and sin. This will make the proof for this case more difficult and may easily cause one to leave a logical gap in the proof. □

Remark. The conclusion derived from the direct proof is stronger: Watson [42, p.43, (2)] at the same time indicates that when ν is an integer, $J_\nu(z)$ and $J_{-\nu}(z)$ are linearly dependent, while the proof using reduction to absurdity has no such function. Thus, sharp tools make good work.

Example 2.2.
If f is integrable and uniformly continuous on $[0,\infty)$, then $|f(t)| \to 0$ as $t \to \infty$.
For the proof of the above statement, read

http://math.stackexchange.com/questions/92105/f-uniformly-continuous-and-int-a-infty-fx-dx-converges-imply-lim-x.

The first two proofs on the web page use reduction to absurdity. The third proof is the most direct one because it does not use reduction to absurdity. However, the discussion given in the the last two lines of the third proof should have divided into two cases: Case $|f(x_0)| < \varepsilon$ and Case $|f(x_0)| \geq \varepsilon$. Furthermore, the third proof reveals the most amount of related information with the least effort. The last two proofs can be generalized to the multidimensional case.

3 The reduced length of an argument that leads to a contradiction

Suppose we try to use the method of reduction to absurdity to prove $A \Rightarrow B$. First, we assume B is false. From this assumption we derive statements $A_1, A_2, \ldots, A_m$ in succession. Suppose we obtain a contradiction when we reach A_m. Then we should go back to check if any A_i can be derived without assuming that B is false. If the answer is yes, we remove it from our list. Finally, we acquire a sublist of statements $B_1, B_2, \ldots, B_n$ which are true only if B is false. We call n the reduced length of the argument that leads to a contradiction. The reduced length separates what we know from what we plan to investigate; without counting the length of direct part of the argument, we use the concept of reduced length to measure the degree of indirectness of an argument that uses reduction to absurdity. A shorter length gives a more straightforward argument. If ($B_1 \Rightarrow B_2 \Rightarrow \ldots \Rightarrow B_n$) and (we obtain a contradiction when we reach B_n), then we know every B_i ($i = 1, \ldots, n$) is false. For example, the argument using reduction to absurdity given in the proof of Ahlfors [1, p.129, Theorem 9] could have included all the statements given in Ahlfors [1, p.128, l.21–p.129, l.10]. Since all these statements are true without assuming that the conclusion of Ahlfors [1, p.129, Theorem 9] is false, we should not put them into the argument for reduction to absurdity even though they are specifically designed for proving Ahlfors [1, p.129, Theorem 9]. Obviously, the reduced length of the proof of Ahlfors [1, p.129, Theorem 9] is shorter than that of the proof of Saks [35, p.144, Theorem 6.1]. This is because Ahlfors adopts a powerful tool (Ahlfors [1, p.128, l.-11–l.-9]) to reach a contradiction. After he uses the tool for the first time, he concludes that $z = a$ is not an essential singularity of $f(z) - A$. After he uses the tool for the second time, he almost reaches a contradiction.

4 The total length of an argument using reduction to absurdity

The total length measures how straightforward the argument is. When we consider the total length of an argument using reduction to absurdity, we have to trace all the theorems we have employed back to axioms. Both the proof given in Dugundji [10, p.362, l.-2–p.363, l.15] and the proof given in Munkres [26, p.391, l.7–l.17] use reduction to absurdity and have about the same length. Does this mean that they are equally straightforward? The answer is no because the length considered aboved is a part of the total length. The total proof length of the Jordan curve theorem is the proof's length plus the sum of the proof lengths of all the theorems used to prove the Jordan curve theorem. Thus, Dugundji [10, p.361, l.6–p.363, l.15] gives the former proof's total length, while Munkres [26, p.385, l.12–p.391, l.17] gives the latter proof's total length. By comparing the two proofs' total lengths, we conclude that the former proof is more straightforward than the latter one.

5 Counting Embedded levels of reduction to absurdity

5.1 n-level reduction to absurdity

If there is no reduction to absurdity embedded in a reduction to absurdity, we say the reduction to absurdity is a 1-level reduction to absurdity. Assume that an $(n-1)$-level reduction to absurdity is defined. The proof of Theorem A begins with the assumption that the conclusion of Theorem A is false; this assumption leads to a contradiction. Suppose during the proof we use Theorem B whose proof requires the use of an $(n-1)$-level reduction to absurdity. Then we say that the proof of Theorem A uses an n-level reduction to absurdity.

Example 5.1.

Let S_n be the statement given in Zygmund [43, vol.1, p.62, (9.4)]. It is unnecessary to use reduction to absurdity when we prove the case $n = 1$.According to the scheme of its proof, we use reduction to absurdity to prove the validity of S_{n+1}. We first assume that S_{n+1} is false. The assumption will lead to a statement contradictory to S_n (Zygmund [43, vol.1, p.62, l.11]). Therefore, S_{n+1} is true. Using the same method, we can prove the validity of S_n. In total, we use n embedded levels of reduction to absurdity to prove the validity of S_{n+1}.

5.2 $\aleph_0$-level reduction to absurdity

Let us compare the proof of Dugundji [10, p.362, Theorem 5.4] with the proof of Munkres [26, p.390, Theorem 63.4]. The proof of Dugundji [10, p.362, Theorem 5.4] uses reduction to absurdity once (Dugundji [10, p.362, l.-2–p.363, l.14]). Although the proof of "the $\Leftarrow$ part" of Dugundji [10, p.359, Proposition 4.1.(1)] uses reduction to absurdity, Dugundji [10, p.362, l.7] only uses "the $\Rightarrow$ part" of Dugundji [10, p.359, Proposition 4.1.(1)].

In contrast, the reduction to absurdity used in the proof of Munkres [26, p.390, Theorem 63.4] starts at Munkres [26, p.391, l.7] and ends at Munkres [26, p.391, l.17]. In Munkres [26, p.391, l.10], Munkres uses Munkres [26, p.389, Theorem 63.2] twice (once for C_1 and once for C_2). The first proof of Munkres [26, p.389, Theorem 63.2] uses Borsuk's lemma (Munkres [26, p.382, Lemma 62.2]). The proof of Munkres [26, p.382, Lemma 62.2] involves 2-level reduction to absurdity. The second proof of Munkres [26, p.389, Theorem 63.2] is divided into two parts: Munkres [26, p.389, l.15–p.390, l.6] and Munkres [26, p.390, l.7–l.26]. Both parts use reduction to absurdity. The induction step in part 2 uses part 1, so the reduction to absurdity for case n is embedded in the reduction to absurdity for case $n+1$. Thus, part 2 of the second proof of Munkres [26, p.389, Theorem 63.2] uses $\aleph_0$-level reduction to absurdity. $\aleph_0$-level reduction to absurdity is the most twisted argument against intuition in logic, we should avoid using this type of argument whenever possible. Because the proof of Dugundji [10, p.362, Theorem 5.4] uses reduction to absurdity fewer times than does the proof of Munkres [26, p.390, Theorem 63.4], the former proof is more straightforward than the latter one.

6 How we shorten the reduced length for an argument using reduction to absurdity

6.1 Understanding the intrinsic nature of reduction to absurdity

The reason we are obliged to use the method of reduction to absurdity is that our resources are limited or that we fail to fully take advantage of available resources. Once our resources are amply supplied or we take full advantage of available resources, we may avoid using the method of reduction to absurdity in our proof. For example, the proof of Birkhoff–Rota [5, p.26, Theorem 7] uses reduction to absurdity, while the proof of Grönwall's inequality [Hartman [19, p.24, Theorem 1.1]] need not use reduction to absurdity. Thus, the method of reduction to absurdity can be considered a detour approach when a direct access is not available. If we use a more elementary method in proving a theorem, the proof will become more straightforward.

Example 6.1. (Uniqueness of Laurent series)

Watson–Whittaker [41, p.100, l.4–l.-10] proves the uniqueness of Laurent series without using reduction to absurdity, while Cauchy's proof given in Watson [42, p.372, l.7–l.10] uses reduction to absurdity. The former proof requires the specific values of a given function f, while the latter proof can be applied to any meromorphic function. In other words, the former proof requires more specific information, while the proof can be applied to the general case.

Example 6.2. (The Jordan curve theorem: the complement of a Jordan curve has at least two components)

The proof of Ahlfors [1, p.118, Exercise 3] does not use the method of reduction to absurdity. In contrast, the proof given in Munkres [26, p.391, l.1–l.6] uses the Jordan separation theorem Munkres [26, p.379, Theorem 61.3] whose proof uses a nontrivial (Munkres [26, p.379, l.-4]) reduction to absurdity. If the Jordan curve is a closed polygon, we can even use elementary geometry to prove this theorem (Saks [35, chap. IV, §11]). Ahlfors' proof is the best because it shows that the concept of winding number is the key to proving the theorem.

We should state our argument using reduction to absurdity as positively as possible. This means that we should perform the following process: First, we should try to make our argument as straightforward as possible. Second, we should identify the statements that are true without assuming that the conclusion of the theorem is false. Third, we should remove these statements from the argument using reduction to absurdity.

Example 6.3.

In the proof of Narasimhan [27, p.27, Theorem 6], we use reduction to absurdity to show $r_0 = r$ (Narasimhan [27, p.27, l.20]). The proof seems quite long (Narasimhan [27, p.27, l.-15–l.-1]). However, we may summarize most of the argument in a positive manner as the following lemma:

Lemma. Let $D = D(a,r)$ and $f \in H(D)$. If f has a primitive on $D(a,q)$, where $q < r$, then there exists an $e > 0$ such that f has a primitive on $D(a,q+e)$.

We may prove this lemma without using reduction to absurdity and without the need to assume $r_0 < r$. Now we assume $r_0 < r$. By the above lemma, f has a primitive on $D(a,q+e)$. This contradicts the definition of r_0 (Narasimhan [27, p.27, l.14]). Thus, the reduced length of the argument using reduction to absurdity is 2.

In order to shorten the length of argument using reduction to absurdity, we should use straightforward reasonings to approach our solution as closely as possible by thoroughly exploring its properties and fully taking advantage of available resources.

Example 6.4.

The proof of Bellman [4, p.10, Exercise 3] uses reduction to absurdity, while the proof of Bellman [4, p.10, Theorem 2] does not use reduction to absurdity. We note that the uniqueness theorem quoted in Bellman [4, p.10, l.21] is a *general* property of solutions, while the property given in Bellman [4, p.10, (4)(a)] is an extra property possessed *olny* by $|Y|$, the determinant of the *specific* solution Y. We also note that Bellman [4, p.10, (3)] has already been the *perfect* answer for $|Y|$.

Sometimes we are forced to use reduction to absurdity simply because we adopt an artificial definition. If we switch to a natural definition, we may avoid using reduction to absurity.

Example 6.5. (How an artifical definition of rigidity necessitates the use of reduction to absurdity in the proof of Chasles' theorem)

All of the following three articles provide the proofs of Chasles' theorem:

Article 1: The Instantaneous Motion of a Rigid Body, Dunham Jackson, The American Mathematical Monthly, Vol. 49, No. 10 (Dec., 1942), pp. 661-667,

Article 2: `https://www.control.isy.liu.se/student/graduate/DynVis/Lectures/le1Chasles.pdf`,

Article 3: `http://www.seas.upenn.edu/~meam520/notes02/EulerChasles4.pdf`.

Here we use only Article 1 and Article 2; Article 3 is used as a side reference because it contains misprints.

Article 1 emphasizes the process of motion, treats a motion as a velocity vector field [Article 1, p.662, l.9], and gives an artificial definition of rigidity [Article 1, p.662, l.15–l.16]. It takes VI intuitive steps [Article 1, §3, I–VI] to prove Chasles' theorem, so we will not get lost in the proof.

In constrast, Article 2 emphasizes the result of motion [the change from the initial position to the final position; the process in between is treated as a black box], treats a motion as a translation plus a bijective linear transformation, and give a natural definition of rigidity (rigid = distance-preserving [Halmos [18, p.142, Theorem]]).

The method of reduction to absurdity is used in the following places in Article 1:

1. P cannot have $\cdots$ [p.664, l.17].
2. this is impossible [p.664, l.19–l.20].
3. repetition of the argument $\cdots$ [p.664, l.23].
4. this velocity can have no $\cdots$ [p.664, l.-15].
5. Since ϕ_1 must $\cdots$ [p.665, l.3].
6. The velocity of P_1' must $\cdots$ [p.665, l.6].
7. and to P_1P_1' $\cdots$ [p.665, l.6–l.7].
8. extend the conclusion to points $\cdots$ [p.665, l.9].
9. P_2 passes through O [p.665, l.13–l.14].
10. the velocities of all points of p_2 $\cdots$ [p.665, l.14].
11. this is impossible [p.665, l.17].

Almost every one of the above places that we use reduction to absurdity is because of adopting the artifical definition of rigidity. In contrast, Article 2 uses the natural definition (rigid = distance-preserving) instead, so it does not use reduction to absurdity.

(Article 1, §3, IV; p.667, l.11–l.13) or Marion–Thornton [24, p.340, (9.39); p.412, l.11–l.16] can be used to explain why a general motion of rigid body can be represented as (1) in Article 2.

A proof using reduction to absurdity often provides only a local view and a shallow level of meaning. Only through a straightforward proof may we have a global view, get to a deep level of meaning, and thus hit the heart of the matter.

Example 6.6. (Uniqueness of the solution to the Dirichlet problem)

Both Wangsness [40, p.87, l.1–l.10] and Wangsness [40, p.172, l.1–l.17] prove the following theorem: Let ϕ be a solution of the Laplace equation. If $\phi \equiv$ constant on ∂V, then $\phi \equiv$ constant on V.
The former proof uses reduction to absurdity, while the latter proof does not.

6.1.1 A physical proof is usually more direct than a geometric proof

Example 6.7. (The geometric criterion for integrability of the Pfaffian differential equation)

For Sneddon [36, p.35, Theorem 8], Carathéodory's thermodynamic proof is more direct than Born's geometric proof because the latter proof uses reduction to absurdity in Sneddon [36, p.38, l.-16–l.-9]. From the similarity between the path given in Sneddon [36, p.36, Fig. 11] and the integral path of Reif [31, p.160, (5·4·2)], we see that Carathéodory's idea originates from solving Reif [31, p.160, (5·4·1)]. His proof is closely related to the measurement of entropy using a quasi-static process. It would be difficult to understand the essence of Carathéodory's proof if one fail to know its physical meaning. In contrast, Born's proof involves only the geometric shape of solutions of the Pfaffian differential equation. One cannot use Born's method to measure entropies.
Remark. Continuously deforming the cylinder [Sneddon [36, p.38, l.-15–l.-14]] refers to reducing the cross section area of σ to 0. The band of accessible points [Sneddon [36, p.38, l.-13–l.-12]] refers to the segment IG_0.

6.2 If the original proof is essentially straightford, we should avoid giving it an outlook of a reduction to absurdity

We should not make things unnecessarily complicated.

Example 6.8.

The proof given in Ince [21, p.212, l.10–p.213, l.3] is straightforward if we eliminate the first four wards and the last sentence. Coddington–Levinson [7, p.292, l.-7] follows Ince's style by using an unnecessary reduction to absurdity. The proof given in Ince [21, p.213, l.-7–-4] does not use reduction to absurdity. However, Coddington–Levinson [7, p.293, l.-7] gives this argument of Ince's the unnecessary outlook of reduction to absurdity again.
Remark. Although the equality given in Ince [21, p.212, l.4] is the same as that given in Coddington–Levinson [7, p.288, (2.5)], the latter form may make it easier to recognize the structural relationship between adjoint boundary-value problems.

6.3 How we take advantage of proof by contradiction

6.3.1 Choosing the method that can eliminate most impossibilities

Example 6.9.

Since A_1B_1 does not intersect A_2B_2, $\alpha \neq 0$ [Bell [2, p.165, l.3]]. Since A_3B_3 does not intersect $A_1B_1 \cup$

$A_2B_2, 0 \neq \beta \neq \alpha$ [Bell [2, p.165, l.3]].
If no two lines given in Bell [2, p.164, l.-3] are parallel and we can follow the procure given in Bell [2, p.164, l.-3–p.165, l.12] to produce a conicoid, then the equation given in Bell [2, p.165, l.12] represents a hyperbolic paraboloid [Bell [2, p.165, l.14]].

Proof 1. (Choose the method that can eliminate most impossibilities)
Since no two lines of the three lines given in Bell [2, p.164, l.-3] are parallel, $l \neq 0, m \neq 0$.
$\Delta = \alpha^2\beta^2 m^2 l^2(\alpha-\beta)^2 \neq 0$. By the Δ-column of Fine–Thompson [14, p.232, Table], we have either case (a) or case (c). Furthermore, $D = 0$, so we have case (c) [Fine–Thompson [14, p.232, Table]]. □

Proof 2. (Find out the most distinguishable difference between two options) [Bell [2, p.163, l.-3–p.164, l.-4]] proves that if three non-intersecting lines are not parallel to the same plane, then any conicoid passing through these three lines is a hyperboloid of one sheet. In fact, the converse is also true. Namely,

Lemma. Any three generators of the same system of a one-sheet hyperboloid cannot be parallel to the same plane.

Proof of the Lemma. The direction-cosines of a λ-generator are proportional to
$(\frac{1}{a}, \frac{-\lambda}{b}, \frac{1}{c}) \times (\frac{1}{a}, \frac{1}{b\lambda}, -\frac{1}{c}) = (\frac{1}{bc}(\lambda - \frac{1}{\lambda}), \frac{-2}{ac}, \frac{1}{ab}(\lambda + \frac{1}{\lambda}))$.
$$\begin{vmatrix} \frac{1}{bc}(\lambda_1 - \frac{1}{\lambda_1}) & \frac{-2}{ac} & \frac{1}{ab}(\lambda_1 + \frac{1}{\lambda_1}) \\ \frac{1}{bc}(\lambda_2 - \frac{1}{\lambda_2}) & \frac{-2}{ac} & \frac{1}{ab}(\lambda_2 + \frac{1}{\lambda_2}) \\ \frac{1}{bc}(\lambda_3 - \frac{1}{\lambda_3}) & \frac{-2}{ac} & \frac{1}{ab}(\lambda_3 + \frac{1}{\lambda_3}) \end{vmatrix} = \frac{-4}{a^2b^2c^2}(\lambda_2-\lambda_1)(\lambda_3-\lambda_2)(\lambda_1-\lambda_3) \neq 0.$$ □

Since the conicoid passes through three non-intersecting lines which are not parallel to the same plane [Bell [2, p.163, l.-3–l.-2]], by the Lemma, the conicoid cannot be a one-sheet hyperboloid. Since no two of the three given lines are parallel, the cases (d) and (e) given in Fine–Thompson [14, p.232, Table] are impossible. □

Proof 3. (Use the main property of the target of proof) A hyperbolic paraboloid is our target of proof and the statement given in Bell [2, p.150, l.2–l.3] is its main property. The direction cosines of the λ-generator $z = \frac{-\beta m y}{\lambda}, lx(\alpha-\beta) - \beta m y = \lambda$ are proportional to
$(0, \frac{\beta m}{\lambda}, 1) \times (l(\alpha-\beta), -\beta m, 0) = (\beta m, l(\alpha-\beta), -\frac{\beta m}{\lambda} l(\alpha-\beta))$. Thus, all the λ-generators are parallel to the plane $lx(\alpha-\beta) - \beta m y = 0$. By the Lemma in Proof 2, the conicoid cannot be a one-sheet hyperboloid. □

6.3.2 Sometimes a contradicton surely occurs on an assigned path

Let Theorem $A = [z \in S \Rightarrow (z \text{ has the property } T)]$ and Theorem $B = [z \notin S \Rightarrow (z \text{ does not the property } T)]$. If Theorem B is true and we have proved that Theorem A is true, then we can prove Theorem B by assuming $z \notin S$ and following the proof of Theorem A. At a certain stage, the proof chain will surely break. Thus, we obtain a contradiction.

Example 6.10. (Line through two given points [Fine–Thompson [14, p.12, l.-4–p.13, l.3]])
Let $z = (x, y)$, $S = P'P''$, and $(z \text{ has the property } T) = [(x, y) \text{ satisfies Fine–Thompson [14, p.12, (1)]}]$.

Remark 1. We may use the parametric representation of a line to avoid using reduction to absurdity, but the above proof is a sure bet to obtain a proof.

Remark 2. Other examples: Fine–Thompson [14, p.52, l.−2–p.53, l.3; p.71, l.−3–p.72, l.2; p.101, l.−7–l.−6].

6.4 Selecting an appropriate target for contradiction

6.4.1 We may replace a distant target with a nearby target by tracing the source of contradiction

Example 6.11. (The expectation of the number of steps till one's ruin)

We want to prove $v(z) = \infty (z \geq 1)$ [Borovkov [6, p.63, l.−3–l.−1]].

Proof. $v(2) = 2v(1) - 2$ [Borovkov [6, p.64, l.4]].
$v(z) = zv(1)$ [by induction].
Thus, $v(1) - 1 = v(1)$. □

Remark. The proof given in Borovkov [6, p.64, l.1–l.14] chooses $v(z) < \infty$ as its target for contradiction, while the above proof chooses $v(1) < \infty$ as its target.

6.4.2 Selecting the target that is most closely related to the theme of the theorem

For an argument using reduction to absurdity, if the target we select is more simply and closely related to the theme of a theorem and our approach is aimed more directly toward our goal, then we will reach a contradiction faster.

Example 6.12. (The set $[0,1]$ is not countable)

The proof of Rudin [33, p.36, Theorem 2.43] uses the concept of compactness. The proof of Royden [32, p.56, Corollary 4] uses the concept of infimum; see the definition of outer measure (Royden [32, p.54, l.9]). In contrast, the proof of Pervin [29, p.20, Theorem 2.1.4] uses a more directly related concept: countability. Consequently, the third proof is the simplest.

Example 6.13.

If we assume that the Jordan curve separates E^2 into more than two components, the assumption will lead to a conclusion which contradicts Dugundji [10, p.362, Theorem 5.2] if we follow the argument of Dugundji [10, p.362, Theorem 5.4], and will lead to another conclusion which contradicts Munkres [26, p.385, Theorem 63.1(c)] if we follow the argument of Munkres [26, p.390, Theorem 63.4]. Dugundji [10, p.362, Theorem 5.2] involves the simple concept "x, y lie in distinct component of $\mathscr{C}J$" (Dugundji [10, p.361, l.−1]), while Munkres [26, p.385, Theorem 63.1(c)] involves the complicated concept of the subgroups of a fundamental group. The former concept has a close relationship with the Jordan curve J, while the latter concept has an indirect relationship with the Jordan curve. Thus, in terms of relationships, Munkres' proof also takes a greater length to reach the contradiction than Dugundji's.

Example 6.14. (If a classification method meets our needs, it will help us reach a contradiction fast)
If $\Delta = 0$ and $|b_1, c_2, d_3| \neq 0$, then the three planes are parallel to one line [Bell [2, p.50, l.8–l.10]].

Proof. The ways that three planes may intersect can be divided into 8 cases:

Case 1: All planes are parallel and distinct.

Case 2: Two planes are coincident, and the third is parallel.

Case 3: All three planes are coincident.

Case 4: Two planes are coincident, and the third cuts the others.

Case 5: Two parallel planes intersect a third plane.

Case 6: Normals are coplanar, planes intersect in pairs, and the intersecting lines are different.

Case 7: Three planes intersect in one line.

Case 8: Three planes intersect at one point.

I The classification method based on the collinearity of normals:

(1) All three normals are collinear: Cases 1, 2 & 3.
(2) Only two normals are collinear: Cases 4 & 5.
(3) none of the normals are collinear: Cases 6, 7, & 8.

II The classification method based on the system's consistency:

(1) Consistent systems: Cases 8, 7, 4, & 3.
(2) Inconsistent systems: Cases 6, 5, 1, & 2.

Since the system is inconsistent [Bell [2, p.49, (4)]], the second classification method meets our needs. If we adopt it instead of the first one, we will eliminate the impossible cases faster. □

Example 6.15.
On a hyperboloid of one sheet, no two generators of the same system intersect [Bell [2, p.155, l.1]]. Bell [2, p.155, l.2–l.5] gives the first proof and Bell [2, p.155, l.6–l.10] gives the second proof. The first proof uses the solution of simultaneous equations as the target for contradition and the solution is the intersecton of the two generators. In contrast, the second proof uses three coplanar generators as the target for contradiction. The target is more removed than that of the first proof in connection with the theme of the theorem. Consequently, the second proof is more complicated than the first one.
Remark. The argument in the second proof requires clarification. We assume that the two generators r,s of the λ-system intersect at R. By Bell [2, p.153, l.-11–l.-9], at most two generators can pass through R, so in the μ-system we may find a generator t not passing through R. By Bell [2, p.155, l.11–l.12], t must intersect r at P and intersect s at Q.

Suppose one wants to go from address A to address B by bus. One will check the network of bus lines. The denser the network, the easier for one to reach B from A. This is because one can find a closer bus stop near A or B. Similarly, building a dense flow chart of theorems for a mathematical subject helps prove Theorem A using reduction to absurdity. This is because finding a theorem with 90% similarity to Theorem

A becomes a simple task if we use the table of contents or the method of organization as our flow chart. The incompatibility of a false assumption and the limited differences between two theorems should help us quickly obtain a contradiction by comparison.

Example 6.16.
A fourth line, D, which does not meet A, B, and C, meets the conicoid in general in two points P and Q, and the generators of the μ-system through P and Q are the only lines which intersect the four given lines A,B,C,D [Bell [2, p.165, l.−15–l.−12]].

Proof. If there were three lines and each of them met A,B,C,D, then, by the statement given in Bell [2, p.165, l.−8–l.−6], A,B,C,D would be four generators of a conicoid. This contradicts the fact that D meets the conicoid in two points P and Q [Bell [2, p.165, l.−15–l.−14]]. □

Remark. In this example, we use Bell [2], a rich and well-organized book in analytic geometry, as our flow chart of theorems.

6.4.3 Selecting a target that is as basic as possible

Using reduction to absurdity to prove $A \Rightarrow B$, we first assume B is false. This assumption will contradict some fact S. If S is an advanced theorem whose proof goes through many logical steps, it will be difficult to determine which step is the main factor causing the contradiction. In this sense, the statement targeted for contradiction should be as basic as possible. For example, if both "S and T" and S are targets that can lead to a contradiction, then we choose S as our target for contradiction.

Example 6.17.
Compare the proof of Rudin [33, p.208, Theorem 9.34] with that in Courant–John [9, vol.2, pp.36–37, §1.4.d].

Example 6.18.
Suppose we want to prove that $\{y \in Y : f_0(y) = f_1(y)\}$ is closed (Massey [25, p.152, l.−16]). We may shorten the proof length by using the general property of a Hausdorff space (Dugundji [10, p.138, 1.2(4)]) rather than the special property of a covering space (Massey [25, p.152, l.−12–l.−6]). The former approach reduces the complexity of the setting and avoids a logical detour.

6.4.4 Selecting a target that is as weak as possible

Using reduction to absurdity to prove $A \Rightarrow B$, we first assume B is false. This assumption will contradict some fact S. If we try to shorten the reduced length of reduction to absurdity, we should select the weakest target. That is, S should be a theorem with minimum conditions. This is because a target requiring fewer conditions to characterize its features allows us to focus on our goal more easily and effectively. This statement can be illustrated using the Venn diagram in set theory: set B contains set A if and only if property B is weaker than property A; suppose $A \Rightarrow B$ is true and we choose B as the target for contradiction, we must prove the statement "*not B*" in order to reach a contradiction; thus, we must consider the complement of B, which is smaller than the complement of A in the Venn diagram; having a smaller scope for consideration is the reason that choosing a weaker target for contradiction makes it easier to obtain a contradiction than choosing a stronger one.

In the following three examples we try to prove if $t \in G$, then $\{2\omega_1, 2\omega_2\}$ is a reduced pair of fundamental periods (González [16, p.370, l.-9–l.-8; p.369, Definition 5.4]).

Example 6.19. (Proof without using reduction to absurdity)

Proof. In order to find a reduced pair of fundamental periods, we must consider periods of the general form: $\omega = 2m\omega_1 + 2n\omega_2$. We try to find a reduced pair among the choices $\{2\omega_1, 2\omega\}$ because
$[|\tau| \geq 1 \Rightarrow (\omega_1$is a nonzero period with smallest absolute value$)]$.
Consider the following cases:

(a) $n = 0$: $Im(\omega/\omega_1) = 0$ (González [16, p.370, l.-7–l.-5]).

(b) $n \notin \{0, 1, -1\}$: $|\omega| > |\omega_2|$ (González [16, p.371, l.16–l.22]).

(c) $n = -1$ and $m = 0$: $Im(\omega/\omega_1) < 0$ (González [16, p.371, l.5–l.6]).

(d) $n = -1$ and $m \neq 0$: $Im(\omega/\omega_1) < 0$ (González [16, p.371, l.14–l.15]).

(e) $n = 1$ and $m = 0$: $max(|\omega_1|, |\omega_2|) < |\omega|$ (González [16, p.371, l.1–l.5]).

(f) $n = 1$ and $m \neq 0$: $|\omega| = |\omega_2|$ (González [16, p.371, l.7–l.11]).

Since we are looking for a reduced pair among the choices $\{2\omega_1, 2\omega\}$, we must discard the cases (a), (b), (c), and (d) from our consideration. By (e) and (f), $\{2\omega_1, 2\omega_2\}$ is a reduced pair of fundamental periods. □

Remark. In González [16, p.371, l.5], González claims that $\{2\omega_1, 2\omega_2\}$ is a reduced pair. This statement is false because he proves $|\omega| \geq |\omega_2|$ only in the case $n = \pm 1$ (González [16, p.371, l.1]) but fails to prove $|\omega| \geq |\omega_2|$ in other cases when he makes the claim. Similarly, the statement given in González [16, p.371, l.11–l.12] is incorrect.

Example 6.20. (Using reduction to absurdity with a strong target for contradiction)

Proof. Assume that the pair $\{2\omega_1, 2\omega_2\}$ is not reduced.
Since $[|\tau| \geq 1 \Rightarrow (\omega_1$is a nonzero period with smallest absolute value$)]$,
there exists a reduced pair $\{2\omega_1, 2\omega\}$, where
$\omega = 2m\omega_1 + 2n\omega_2 \quad (*)$.
In the following, we use $(*)$ as our target for contradiction.
By González [16, p.369, Definition 5.4], we have
$Im(\omega/\omega_1) > 0$ and $|\omega| < |\omega_2| \quad (**)$.
Consider the following cases:

(a) $n = 0$: $Im(\omega/\omega_1) = 0$ (González [16, p.370, l.-7–l.-5]).
By González [16, p.369, Definition 5.4], $\{2\omega_1, 2\omega\}$ cannot be a reduced pair. This contradicts $(*)$.

(b) $n \notin \{0, 1, -1\}$: $|\omega| > |\omega_2|$ (González [16, p.371, l.16–l.22]).
By González [16, p.369, Definition 5.4], $\{2\omega_1, 2\omega\}$ cannot be a reduced pair. This contradicts $(*)$.

(c) $n=-1$ and $m=0$: $Im(\omega/\omega_1)<0$ (González [16, p.371, l.5–l.6]).
By González [16, p.369, Definition 5.4], $\{2\omega_1, 2\omega\}$ cannot be a reduced pair. This contradicts $(*)$.

(d) $n=-1$ and $m\neq 0$: $Im(\omega/\omega_1)<0$ (González [16, p.371, l.14–l.15]).
By González [16, p.369, Definition 5.4], $\{2\omega_1, 2\omega\}$ cannot be a reduced pair. This contradicts $(*)$.

(e) $n=1$ and $m=0$: $max(|\omega_1|, |\omega_2|)<|\omega|$ (González [16, p.371, l.1–l.5]).

(f) $n=1$ and $m\neq 0$: $|\omega|=|\omega_2|$ (González [16, p.371, l.7–l.11]).

By (e) and (f), $\{2\omega_1, 2\omega_2\}$ is a reduced pair. This contradicts our assumption: $\{2\omega_1, 2\omega_2\}$ is not reduced. Thus, our assumption is false. □

Example 6.21. (Using reduction to absurdity with a weak target for contradiction)

Proof. $(*) \Rightarrow (**)$. Thus, $(*)$ is a stronger target for contradiction. Now we use $(**)$ as our target for contradiction instead. The proof is the same as Example 6.20 except that we must replace each red-colored line "By González [16, p.369, Definition 5.4], $\{2\omega_1, 2\omega\}$ cannot be a reduced pair. This contradicts $(*)$. " with the statement "This contradicts $(**)$." □

Example 6.22. (Using reduction to absurdity with a weak target for contradiction)

Suppose we want to prove the Abel–Ruffini Theorem (Jacobson [23, vol.3, p.104, l.−2–l.−1]). The theorem follows from Jacobson [23, vol.3, p.104, Theorem 7], van der Waerden [37, vol.1, p.149, Theorem] and Edwards [11, p.61, l.17–l.18]. We use the Galois group's solvability as our target for contradiction. Note that the solvability implies the conclusion of Edwards [11, p.91, Galois' Theorem]. If we use the conclusion of Edwards [11, p.91, Galois' Theorem] as our target for contradiction, the proof becomes much more straightforward (Edwards [11, p.91, l.10–l.16]). Indeed, the comparison of natural numbers is much simpler than the determination of the Galois group's solvability. Thus, the reduction of absurdity used to prove the Abel–Ruffini Theorem becomes trivial. Consequently, Galois' proof using reduction to absurdity given in Edwards [11, p.91, l.10–l.16] is more straightforward than Abel's proof using reduction to absurdity given in van der Waerden [37, vol.1, p.176, l.−6–p.177, l.13]. Edwards [11, p.91, Galois' Theorem] allows us to explore more about solvability and obtain a special result under a particular condition. Edwards [11, p.65, l.10–l.14; p.91, l.17–l.23] discusses more advantages about Galois' approach. In the direction of (Not the conclusion of Edwards [11, p.91, Galois' Theorem]) $\Rightarrow$ (Not the hypothesis of Edwards [11, p.91, Galois' Theorem], i.e., the Galois group is not solvable) $\Rightarrow$ (The general equation of degree 5 is not solvable), we pinpoint a reason for contradiction by focusing on fewer conditions.
Remark. The proof of Jacobson [23, vol.3, p.107, Theorem 8] actually provides the method of constructing insolvable polynomials of prime degree with integer coefficients.

Example 6.23. (The Jordan canonical form)

The proof of Nelson [28, Theorem 8] does not use reduction to absurdity. In view of Nelson [28, (9) and the example immediately after the proof of Theorem 8], we see that reducing a matrix to the Jordan canonical form is nothing but a clever method of counting basis elements. Only after the method is well organized may we avoid using reduction to absurdity. Mathematical induction is useful for organizing the proof structure since it can pinpoint the crucial step. In view of the counting method given in Nelson [28, the example immediately after the proof of Theorem 8], we find that Nelson [28, Theorem 8] counts basis elements row by row, while Herstein [20, p.294, Theorem 6.5.1] counts basis elements column by column. Counting row by row is simpler than counting column by column because we can take advantage of the embedded structure

of Nelson [28, (9)]. More specifically, Nelson [28, Theorem 8] counts basis elements in $\mathscr{N}N_i/\mathscr{N}N_{i-1}$, while Herstein [20, p.295, Lemma 6.5.4] counts basis elements in V/V_1. The number of elements in latter basis is much larger than that in the former basis, so it is more difficult for the reduction to absurdity used in Herstein [20, p.295, Lemma 6.5.4] to reach a contradiction. In fact, the proof of Herstein [20, p.295, Lemma 6.5.4] assumes $V \neq V_1 + W$ and tries to prove that $V/(V_1+W)$ has an empty basis by using the fact given in Herstein [20, p.295, l.15–l.16]. We should move Herstein [20, p.295, l.6–l.17] outside of the argument of reduction to absurdity used in the proof of Herstein [20, p.295, Lemma 6.5.4] because Herstein [20, p.295, l.15–l.16] is a straightforward consequence of "Herstein [20, p.295, l.4–l.5] and $z \notin V_1 + W$". Only the assumption that W has the largest possible dimension will contradict the fact given in Herstein [20, p.295, l.15–l.16].
Remark. It greatly clarifies the proof structure of the uniqueness part when Herstein separates the case of nilpotent transformations from the general case. Although the proof Nelson [28, Theorem 8] focuses on the essential part of the process for producing the Jordan normal form and provides a straightforward logical structure of the argument, the method of the proof is not as effective and practical as the method given in Jacobson [23, vol.2, chap.III, §11, p.92, l.-10–p.94, l.-4]. This is because in practice it is difficult to obtain $\beta_1, \ldots, \beta_k$ given in the proof of Nelson [28, Theorem 8]. In contrast, Jacobson [23, vol.2, chap.III, §11, p.92, l.-10–p.94, l.-4] theorizes the practical method of finding the classical canonical form of a matrix. Since the design is based on practical calculations, its argument is straightforward. The existence of classical canonical form given in Jacobson [23, vol.2, chap.III, §11] can be generalized to the case when $\mathfrak{R}$ is a $\mathfrak{o}$-module, where $\mathfrak{o}$ is a principal ideal domain (Jacobson [23, vol.2, chap.III, p.86, Theorem 6, p.86, l.-5–l.-3; p.88, Theorem 7]). The uniqueness of classical canonical form given in Jacobson [23, vol.2, chap.III, §11] follows from Jacobson [23, vol.2, p.91, Theorem 11 and Theorem 12]. Thus, for the uniqueness of normal form, the proof given in Nelson [28, Theorem 8] is more concrete than the proof given in Jacobson [23, vol.2, p.91, Theorem 10] because the setting in the former proof is more resourceful.

Example 6.24. (Higher order linear differential equations with constant coefficients)

We want to prove the vectors given in Coddington–Levinson [7, p.89, (6.20)] are linear independent. The proof given in Collatz [8, II§4, sec.12] uses the statement given in Collatz [8, p.97, l.19–l.22] as the target for contradiction. The proof of Pontryagin [30, pp.50–51, Theorem 5] straightens the above argument somewhat by exploring the structure more and by using Pontryagin [30, p.52, (B)] as the target for contradiction. The following method allows us to prove Coddington–Levinson [7, p.89, Theorem 6.5] without using reduction to absurdity
Let V be the space generated by the vectors given in Coddington–Levinson [7, p.89, (6.20)], D be the derivative operator on V, and $V_i = \mathscr{N}(D - \lambda_i I)^{m_i}$.
Suppose $\sum_{i=1}^{s}\sum_{k=1}^{m_i-1} c_{ik}t^k e^{\lambda_i t} = 0$.
Since $\sum_{k=1}^{m_i-1} c_{ik}t^k e^{\lambda_i t} \in V_i$,
$\sum_{k=1}^{m_i-1} c_{ik}t^k e^{\lambda_i t} = 0 (i = 1, \ldots, s)$ (Nelson [28, Theorem 6]).
Fix i. We have $\sum_{k=1}^{m_i-1} c_{ik}t^k = 0$.

From above examples, we see that the use of a weaker target for contradiction allows us to shorten the reduced length of the argument using reduction to absurdity. A stronger target requires more conditions to characterize its features. In contrast, a weaker target requires fewer conditions to characterize its features. If we are able to go to the extreme and choose a target with no conditions, then we will have a proof without using reduction to absurdity.

7 How we shorten the total length for an argument using reduction to absurdity

Suppose there are two proofs for a theorem. Both of them do not use reduction to absurdity. If one proof is long and the other one is short, then the shorter one is more direct. Thus, in order to determine whether an argument is direct, we cannot consider reduction to absurdity alone. At the same time we see that it is more appropriate to consider the total length of an argument using reduction to absurdity rather than its reduced length.

Example 7.1. (The uniqueness of the periodic solution for Mathieu's equation)

Both Guo–Wang [17, p.617, l.-12–p.619, l.13] and Elsevier [13, §32.3, NONPERIODICITY OF SECOND SOLUTION, p.20] prove the following statement:

If $q \neq 0$, then there is at most one periodic solution for Guo–Wang [17, p.610, (1)].

Guo–Wang [17, p.610, (1)] is a special case of Guo–Wang [17, p.614, (1)]. In a more generalised setting, there are fewer conditions that a mathematical term must satisfy. Therefore, it is easier to identify two objects by comparing their properties. Since the proof for uniqueness is an argument designed for a more generalized setting, the more general the case is, the shorter the proof is (Wang [39, §3.1]) even though the more concrete case may have more flexible proofs. If we consider the total lengths of the above two proofs, the length given in Guo–Wang [17, p.617, l.-12–p.619, l.13] should be shorter than the length given in Elsevier [13, §32.3, NONPERIODICITY OF SECOND SOLUTION, p.20].

Remark 1. The existence of a Floquet solution given in Guo–Wang [17, p.614, l.7] can be found in Elsevier [13, §32.3, FLOQUET'S THEOREM, p.21].

Remark 2. If we impose an extra condition "one periodic solution is an even function" on Elsevier [13, §32.3, NONPERIODICITY OF SECOND SOLUTION, p.20], then the proof for this case is shorter than the proof for the case when one periodic solution may be even or odd. This can easily be seen via Venn diagram. Thus, if we use only the method designed for a concrete setting, then a more specialized case may have a shorter proof.

8 The concept of straightforwardness can be used to determined the quality of a proof

Example 8.1. (The Sturm comparison theorem)(Birkhoff–Rota [5, pp.267–268, §10.6])

The reduction to absurdity used in the proof given in Ince [21, §10.31] is contained in Ince [21, p.226, l.12–l.15]. All it does is compare a positive number with 0, so the reduction to absurdity used here is trivial. Now we count the number of times that reduction to absurdity used in the proof of Birkhoff–Rota [5, p.268, Theorem 3].

(1) Prerequisites

  (a) The proof of Birkhoff–Rota [5, p.26, Theorem 7] uses a nontrivial reduction to absurdity given in Birkhoff–Rota [5, p.26, l.-12–l.-3]. The reduction to absurdity is nontrivial because Birkhoff–Rota [5, p.24, Lemma 2] is used in Birkhoff–Rota [5, p.26, l-6].
  (b) The proof of Birkhoff–Rota [5, pp.26–27, Theorem 8] uses Birkhoff–Rota [5, p.26, Theorem 7] twice.
  (c) Birkhoff–Rota [5, p.27, Corollary 1] follows from Birkhoff–Rota [5, pp.26–27, Theorem 8].
  (d) The proof of Birkhoff–Rota [5, p.27, Corollary 2] uses another nontrivial reduction to absurdity given in Birkhoff–Rota [5, p.27, l.-11–l.-5]. The reduction to absurdity is nontrivial because Birkhoff–Rota [5, pp.26–27, Theorem 8] is used in Birkhoff–Rota [5, p.27, l.-7].

(2) (a) Birkhoff–Rota [5, pp.26–27, Theorem 8] is used in Birkhoff–Rota [5, p.268, l.9].
  (b) In order to prove $\theta(x) \equiv \theta_1(x)$ given in Birkhoff–Rota [5, p.268, l.12], we must use Birkhoff–Rota [5, p.27, Corollary 2].
  (c) In order to prove the statement given in Birkhoff–Rota [5, p.268, l.16–l.17], we must use Birkhoff–Rota [5, pp.26–27, Theorem 8, Corollary 1 & Corollary 2].

Remark. The reduction to absurdity used in Ince [21, §10.31] is trivial because the Picone formula reaches deeply into the essence of the problem. In contrast, the proof of Birkhoff–Rota [5, p.268, Theorem 3] uses reduction to absurdity so many times because the proof remains at a shallow level from beginning to end. Birkhoff–Rota [5, pp.266–267, §10.5] is nothing but a complication. Thus, Birkhoff and Rota's proof of the Sturm comparison theorem is far less straightforward than Picone's proof. A more straightforward proof is a better proof. Authors of textbooks in mathematics should choose better proofs to put into their books.

9 Conclusion

The history of mathematics can be considered the history of evolution of the method of reduction to absurdity. In the beginning of solving mathematical problems, many mathematicians make conjectures and try to find their rigorous proofs. If we use reduction to absurdity in a proof, we try to improve the quality of the proof by making it more staightfoward. The inability to find a proof without using reduction to absurdity indicates that there is something deeper needed to explore. Our final goal is to find a proof without using reduction to absurdity. However, making an argument straightforward may take quite a unexpectedly long way. For example, we can use the method of Lebesgue integraion to make some argument in calculus of variation straightforward: using the proof of Rudin [34, p.31, Theorem 1.39(a)] to make the proof of Fomin–Gelfand [15, p.9, Lemma 1] straightforward as given in the proof of Rudin [34, p.31, Theorem 1.39(b)].

Example 9.1. (Even though a theorem involves the axiom of choice, we can still make its reduction-to-absurdity proof straightforward)

Borovkov [6, p.431, l.-6–p.432, l.-3] uses the method of reduction to absurdity to prove

$$\mathcal{B}_n \downarrow \emptyset \Rightarrow \lim_{n\to\infty} P(\mathcal{B}_n) = 0.$$

By Bell [3, p.4, Lemma 9; p.5 Lemma 10], we can prove the Kolmogorov extension theorem [Borovkov [6, p.431, Theorem]] without using reduction to absurdity. Using the concept of compact class is the key to

removing this reduction to absurdity.
Remark. Both Borovkov [6, p.428, Theorem 1] and Bell [3, pp. 1–2, Theorem 2] use the same key idea: the outer measure.

References

[1] Ahlfors, L. V.: *Complex Analysis*, 2nd ed., New York: McGraw-Hill, 1966.

[2] Bell, R. J. T.: *An Elementary Treatise on Coordinate Geometry of Three Dimentions*, 3rd ed., London: Macmillan, 1963.

[3] Bell, J: http://individual.utoronto.ca/jordanbell/notes/kolmogorov.pdf

[4] Bellman, R.: *Stability Theory of Differential Equstions*, New York: McGraw-Hill, 1953.

[5] Birkhoff, G. & Rota, G. C.: *Ordinary Differential Equations*, 3rd ed., New York: John Wiley & Sons, 1978.

[6] Borovkov, A. A.: *Probability Theory* (translated from the Russian by O. Borovkova), Amsterdam: Gordon and Breach Science Publishers, 1998.

[7] Coddington E. A. & Levinson N.: *The Theory of Ordinary Differential Equations*, New York: McGraw-Hill, 1955.

[8] Collatz, L.: *Differential Equations* (translated from German by E. R. Dawson), New York: John Wiley & Sons, 1986.

[9] Courant R. & John, F: *Introduction to Calculus and Analysis*, 2 vols., New York: John Wiley & sons, vol.1 (1965), vol.2 (1974).

[10] Dugundji, J.: *Topology*, Boston: Allyn and Bacon, 1966.

[11] Edwards, H. M.: *Galois Theory*, New York: Springer-Verlag, 1984.

[12] Ellison, W. & F.: *Prime Numbers*, New York: John Wiley & Sons, 1985.

[13] Elsevier: http://www.elsevierdirect.com/companions/9780123846549/Chap_Mathieu.pdf

[14] Fine, H. B. & Thompson H. D.: *Coordinate Geometry*, New York: The Macmillan Company, 1909.

[15] Fomin, S. V. & Gelfand, I. M.: *Calculus of Variations*, translated and edited by R. A. Silverman, Englewood Cliffs, NJ: Prentice-Hall, 1963.

[16] González, M. O.: *Complex Analysis (Selected Topics)*, New York: Marcel Dekker, 1992.

[17] Guo, D. R. & Wang, Z. X.: *Special Functions*, translated from the Chinese by D. R. Guo & X. J. Xio, Singapore: World Scientific, 1989.

[18] Halmos, P. R.: *Finite Dimensional Vector Spaces*, 2nd ed., Princeton: D. Van Nostrand, 1958.

[19] Hartman, P.: *Ordinary Differential Equations*, 2nd ed., Boston: Birkhäuser, 1982.

[20] Herstein, I. N.: *Topics in Algebra*, 2nd ed., Lexington: Xerox College, 1975.

[21] Ince, E. L.: *Ordinary Differential Equations*, New York: Dover, 1956.

[22] Ireland, K. & Rosen, M.: *A Classical Introduction to Modern Number Theory*, 2nd ed., New York: Springer-Verlag, 1990.

[23] Jacobson, N.: *Lectures in Abstract Algebra*, 3 vols., Princeton: Van Nostrand, vol. 1 (1951), vol. 2 (1953), vol.3 (1964).

[24] Marion, J. B. & Thornton, S. T.: *Classical Dynamics of Particles and Systems*, 5th ed., Belmont, C.A.: Brooks/Cole, 2004.

[25] Massey, W. S.:*Algebraic Topology: An Introduction*, New York: Springer-Verlag, 1989.

[26] Munkres J. R.: *Topology*, 2nd ed., Englewood Cliffs, NJ: Prentice-Hall, 2000.

[27] Narasimhan R.: *Complex Analysis in One Variable*, Boston: Birhäuser, 1985.

[28] Nelson, E.: http://www.math.princeton.edu/ nelson/217/jordan.pdf

[29] Pervin, W.J.: *Foundations of General Topology*, New York, Academic Press, 1964.

[30] Pontryagin, L. S.: *Ordinary Differential Equations*, translated from the Russian by Leonas Kacinskas and Walter B. Counts, Reading: Addison-Wesley Publishing Company, 1962.

[31] Reif, F.: *Statistical Thermal Physics*, New York: McGraw-Hill, 1965.

[32] Royden, H. L.: *Real Analysis*, 2nd ed., New York: The Macmillan Company, 1968.

[33] Rudin, W.: *Principles of Mathematical Analysis*, 2nd ed., New York: McGraw- Hill, 1964.

[34] Rudin, W.: *Real and Complex Analysis*, 2nd ed., New York: McGraw-Hill, 1974.

[35] Saks, S. & Zygmund, A.: *Analytic Functions*, translated by E. J. Scott, Monografje Matematyczne, vol. 28, 3rd ed., Warsaw, 1971.

[36] Sneddon, I. N.: *Elements of Partial Differential Equations*, New York: McGraw-Hall, 1957.

[37] van der Waerden, B. L.: *Modern Algebra*, 2 vols, Translated from the German by F. Blum, New York: Ungar, 1949.

[38] Wang, L. C.: http://www.lcwangpress.com/physics/absurd-de.html

[39] Wang, L. C.: http://www.lcwangpress.com/papers/existence.pdf

[40] Wangsness, R. K.: *Electromagnetic Fields*, New York: John Wiley & Sons, 1986.

[41] Watson, G. N. & Whittaker E. T.: *A Course of Modern Analysis*, 4th ed., Cambridge: Cambridge University Press, 1963.

[42] Watson, G. N.: *Theory of Bessel Functions*, 2nd ed., Cambridge: Cambridge University Press, 1966.

[43] Zygmund, A.: *Trigonometric Series*, 2 vols., 2nd ed., Cambridge: Cambridge University Press, 1959.

Mr. Li-Chung Wang is the author of the following website about the philosophy of mechanics:
http://www.lcwangpress.com/physics/main.html.
Address: 7th Floor, #21 Lane 267, Xi-zhou Street, Chungli, Taiwan, ROC.
E-mail:lcwangpress@yahoo.com