Limitations of Spacecraft Redundancy: A Case Study Analysis

Robert P. Ocampo[1]
University of Colorado Boulder, Boulder, CO, 80309

Redundancy can increase spacecraft safety by providing the crew or ground with multiple means of achieving a given function. However, redundancy can also *decrease* spacecraft safety by 1) adding additional failure modes to the system, 2) increasing design "opaqueness", 3) encouraging operational risk, and 4) masking or "normalizing" design flaws. Two Loss of Crew (LOC) events—Soyuz 11 and *Challenger* STS 51-L—are presented as examples of these limitations. Together, these case studies suggest that redundancy is *not necessarily* a fail-safe means of improving spacecraft safety.

I. Introduction

A redundant system is one that can achieve its intended function through multiple independent pathways or elements[1,2]. In crewed spacecraft, redundancy is typically applied to systems that are critical for safety and/or mission success[3,4]. Since no piece of hardware can be made perfectly reliable, redundancy—in theory—allows for the benign (e.g. non-catastrophic) failure of critical elements.

Redundant elements can be 1) similar or dissimilar to each other, 2) activated automatically ("hot spare") or manually ("cold spare"), and 3) located together or separated geographically[5-7]. U.S. spacecraft have employed redundancy on virtually all levels of spacecraft design, from component to subsystem[7,8].

Redundancy has a successful history of precluding critical and catastrophic failures during human spaceflight. A review of NASA mission reports, from Mercury to Space Shuttle, indicates that redundancy has saved the crew or extended the mission over 160 times, or roughly once per flight[9]. The presence of a partially redundant spacecraft (the Lunar Module) during the Apollo 13 emergency notably contributed to the crew's safe return[10].

Despite its frequent success, redundancy has *also* been responsible for a disproportionate share of catastrophic failures: of the four Loss of Crew (LOC) events that have occurred during the space age, *three* can be directly attributed to the failure of a redundant element to perform as expected. Two of these missions—Soyuz 11 and *Challenger* STS-51-L[2] —are discussed below.

II. Case Studies

A. Soyuz 11

In 1971, the crew of Soyuz 11 asphyxiated when one of two redundant pressure relief valves inadvertently opened after the planned separation of the orbital and descent modules. Gunpowder was later discovered in the faulty valve[11], (valve 1) indicating that pyrotechnic separation of the two modules triggered the valve's fatal opening.

At the time of the disaster, Soyuz employed two redundant pressure relief valves in its capsule design (see Fig. 1a). Although only one valve was necessary to equalize cabin pressure during re-entry, Soyuz 11 flew with two pressure relief valves to ensure fresh air was available to the crew in the event the capsule landed in water[11]. Both valves had two shutters—one which operated manually, and one which opened automatically after the deployment of the main parachute. The shutters were placed in series so that capsule depressurization through a given valve would only occur if *both* shutters were open (Fig. 1a).

---

[1] PhD Student, Aerospace Engineering Sciences, 429 UCB, Boulder, CO 80309

[2] Redundancy also played a role in the crash landing of Soyuz 1 in 1967. In this instance, neither the primary nor backup parachute functioned properly, leading to the death of cosmonaut Vladimir Komarov. The fourth LOC event—the loss of the Space Shuttle *Columbia* in 2003—cannot be directly attributed to the failure of a redundant element. However, in the case of *Columbia*, no form of redundancy would have precluded the vehicle's catastrophic loss.

Nominal procedures called for three of the four shutters to be set closed prior to launch: both shutters in valve 1, and the automatic shutter in valve 2 (with the manual shutter in valve 2 left open—see Fig. 1a). However, during valve installation, the two valves were inadvertently switched, leaving all shutters closed except for the manual shutter in valve 1[11] (Fig. 1b).

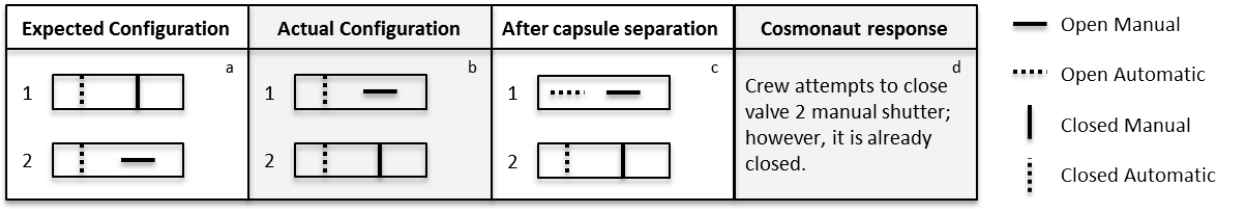

Figure 1. Soyuz 11 expected and actual depressurization valve configuration. *(a) Soyuz 11 pressure relief valves in their expected pre-launch configuration. On valve 1, both manual and automatic shutters were supposed to be closed; on valve 2, only the automatic shutter was supposed to be closed. (b) Due to incorrect valve installation, valves 1 and 2 were switched: valve 2 had both shutters closed, while only the automatic shutter on valve 1 was closed. (c) Capsule separation triggered the automatic shutter in valve 1 to open. This created an open path between the cabin and the external environment, leading to cabin depressurization. (d) Because the crew believed the valves were in their expected configuration, they attempted to close the manual shutter on valve 2. This was the proper response given the system's expected configuration. However, because the valves were switched, this action was ineffective; the crew spent their final minutes trying to close a valve that was already closed.*

During reentry, capsule separation inadvertently triggered the automatic shutter in valve 1 to open, thereby creating an open path between the cabin and the external vacuum. The cosmonauts, *not knowing the two valves had been switched*, assumed the automatic shutter in valve 2 had been breached, and attempted to close it manually[11,12] (Fig. 2d). Given the cosmonaut's knowledge of the *nominal* valve configuration, this was the appropriate response. However, because the two valves had been switched, the crew spent their final seconds trying to manually close a valve *that was already closed*.

B. Challenger (STS-51-L)

The space shuttle *Challenger* disintegrated during launch when heated exhaust gas eroded, then breached both primary and secondary seals in the right Solid Rocket Booster (SRB) aft field joint. The escaping gas destroyed the strut connecting the External Tank (ET) to the SRB, leading to the vehicle's structural breakup and the loss of all seven crewmembers on board[13,14].

Each SRB was comprised of seven segments connected by pins at tang and clevis joints. Two O-rings—one primary, and one secondary—were positioned in series within the joint so as to maintain a gastight seal between segments (Fig. 2a). During ignition, SRB bending (known within NASA as "joint rotation") increased the gap between tang and clevis. During this transient time period, joint sealing was nominally[3] maintained through pressure actuation—a process whereby combustion gases extruded the O-ring in to the gap[13,15,16] (Fig 2b-2c).

On STS-51-L, primary O-ring sealing was precluded for two reasons: 1) a smaller than normal pre-launch gap between tang and clevis, resulting from an "out-of-round" SRB segment, eliminated the requisite spacing between O-ring and upstream wall (Fig. 2d). This prevented combustion gases from getting behind the O-ring and initiating sealing during joint rotation[13] (Fig. 2e-2f). In addition, extreme cold at launch (estimated to be 28 ± 5° F at the joint) significantly compromised the resiliency of the rubber O-ring, thereby further reducing the rate at which the O-ring could expand into its seat[13].

The secondary O-ring proved ineffective as a redundant backup, as joint rotation, coupled with the extreme cold, prevented it from effectively sealing the joint. Aluminum oxides, a by-product of the burned SRB propellant, managed to temporarily seal the joint during liftoff, preventing a launch pad catastrophe; however severe wind shear at T +37 seconds destroyed this temporary seal[13], leading to the eventual aerodynamic destruction of the vehicle.

[3] Neither joint rotation not pressure actuation were anticipated during SRB design. John Q. Miller, chief of the Solid Rocket Motor branch at the time of the Challenger accident went so far as to state that the sealing technique "violate[d] industry and government O-ring application practices[13]." Nevertheless, pressure actuation successfully maintained gastight sealing of the SRB for the first 24 shuttle launches, despite multiple instances of primary and secondary O-ring erosion.

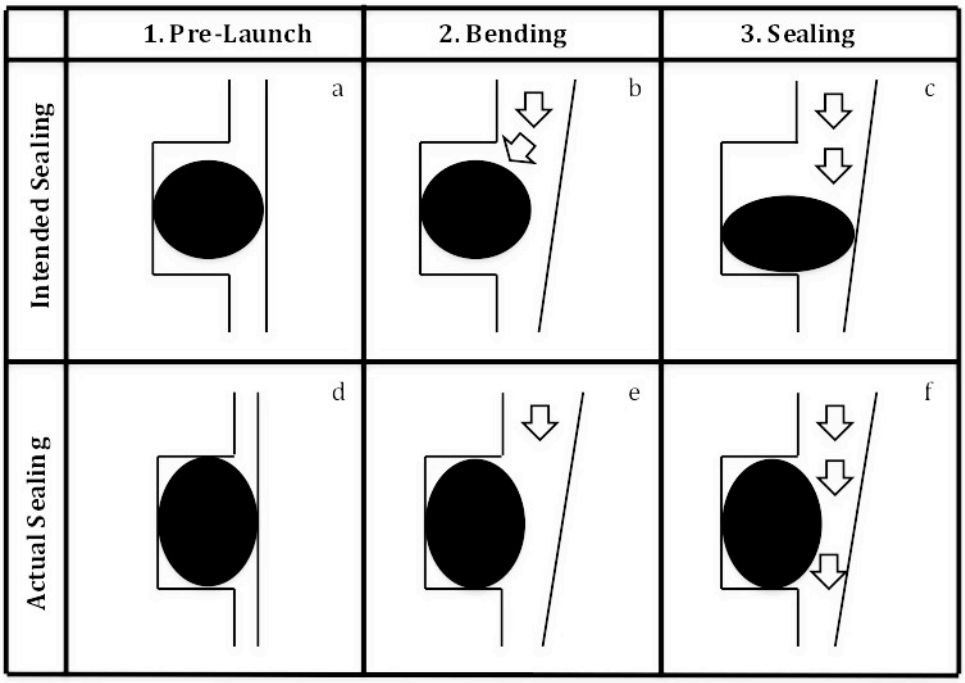

Figure 2. Intended and Actual Sealing Characteristics of SRB O-rings. *Ideally, the O-ring should be seated so that neither the upstream nor downstream wall is touching (a). When joint bending occurs, this configuration allows combustion gases to get behind the O-ring (b), sealing the joint through pressure actuation (c). The leaking joint on STS-51-L was out of round, causing the O-ring to compress and touch the upstream wall (d). Because of this compression, combustion gases could not get behind the O-ring to initiate sealing during joint bending. Furthermore, cold weather at launch reduced O-ring resiliency, thereby allowing the combustion gases to breach the joint.*

III. Discussion

These two case studies suggest redundancy can contribute to catastrophic spacecraft failure in at least four key ways:

A. Redundancy adds potential failure modes to the design

The addition of redundant elements invariably adds complexity to crewed spacecraft[17-21], often in a non-linear fashion[22]. Passive redundant hardware takes up volume and adds mass, while active redundant hardware necessitates separate power, cooling, and data distribution networks[6]. This added complexity introduces novel (and sometimes unexpected) failure modes that can reduce crew safety[6,17-19].

• The Titan III solid rockets, from which the Space Shuttle SRBs were derived, utilized only a single O-ring[13]. To accommodate a *second* O-ring on the Space Shuttle, engineers were forced to extend the SRB tang, thereby increasing joint rotation during launch[13,23]—an effect that actually reduced the sealing effectiveness of the primary O-ring and negated the functionality of the second O-ring.

• The addition of a second depressurization valve on the Soyuz descent capsule was intended to supplement the crew's limited air supply during re-entry and landing; on its own, the descent capsule's small pressurized volume could only support three cosmonauts for a short time[11]. In actuality, the second depressurization valve on Soyuz 11 actually served to increase both the likelihood and rate of catastrophic depressurization[4].

[4] Although conjecture, it is reasonable to assume that had there only been *one* depressurization valve on board Soyuz 11, both manual and automatic shutters would have been set closed prior to launch. In this scenario, even if capsule separation had triggered the opening of the automatic shutter, the manual shutter would have been sufficient to maintain cabin pressurization.

B. Redundancy increases design opacity

The greater the number of parts (redundant or otherwise), the harder it is for engineers to verify or assess the overall state of the system[6,18,21]. Although such opacity does not reduce safety in and of itself, it reduces the *likelihood* that the crew or ground will correctly predict or react to impending failures.

• The presence of a visually identical—but differentially configured—redundant depressurization valve made it nearly impossible for either the ground or flight crew to recognize its improper installation on Soyuz 11. As a result, the cosmonauts were unable to select the correct valve to close during depressurization.

• By 1986, the second O-ring was no longer classified by NASA as a redundant element[13,16,24]. However, due to errors and limitations in NASA's Problem Assessment System, several NASA managers were led to believed that redundancy existed in Challenger's SRB field joint prior to launch[13,25]. According to the Rogers Commission—the committee assigned by the White House to investigate the *Challenger* accident—this mis-categorization of SRB redundancy made "informed decision making…impossible" and may have contributed to the decision to launch *Challenger* in temperatures that fell outside its designed operating range[13].

C. The presence of redundant elements can encourage greater risk taking

In redundant systems, primary elements are often operated at or beyond their nominal operating envelope when backup elements are thought to be in place[6,18,20,26]. Such risk taking can undermine or offset any safety benefits associated with redundancy.

• None of the Soyuz 11 crewmembers wore pressure suits during re-entry[11], despite the depressurization risks associated with a second relief valve. This is not to say that redundancy directly led the cosmonauts to fly without pressure suits, but it suggests confidence in the descent module's design—including its set of redundant depressurization valves. According to Nikolai Kamanin, head of the cosmonaut program at the time of Soyuz 11, "the decompression of the Soyuz spacecraft [was] completely excluded" as a potential failure mode, and that cosmonauts could fly in "shorts" if necessary[11,12].

• Faith in the O-ring's redundancy is thought to have encouraged NASA mangers to launch *Challenger* in record-cold temperatures[13,26]. Although flight[5] and ground tests had shown that O-ring sealing ability decreased exponentially with temperature, the Flight Readiness Review (FRR) conducted prior to STS-51-E concluded that low temperature was an "acceptable risk because of limited exposure and [O-ring] *redundancy* [emphasis added][13,27]. Despite predicted launch temperatures 15° colder than any previous launch, this rationale was neither revisited nor explicitly discussed in the FRR that preceded *Challenger* STS-51-L.

D. Redundancy can discourage further optimization of individual elements

Redundancy can mask or "normalize"[25] design flaws that may have otherwise been rectified in single-element (i.e. non-redundant) hardware. These design flaws, if left unresolved, can then propagate in unexpected and (potentially) catastrophic fashion.

• Engineers first recognized the SRB joints were not sealing as designed in 1981, after the flight of STS-2[13]. By 1985, O-ring erosion had become so endemic (occurring at least 38 times in the 24 flights preceding *Challenger*[13]) that Morton Thiokol, the subcontractor responsible for the SRB, began submitting new seal concepts to NASA for review[13]. Yet despite the prevalence of O-ring erosion—a phenomenon serious enough to warrant SRB redesign—STS-51-L was certified to fly under the assumption that any effects of primary O-ring erosion would be mitigated by "limited exposure and [O-ring] *redundancy* [emphasis added]."[13,27]

[5] The SRBs on STS-51-C, launched a year prior to Challenger in then record-cold temperatures, experienced the greatest O-ring damage to date[13]. The weather for *Challenger* STS-51-L was *15°F colder* than STS-51-C.

	Soyuz 11	Challenger STS-51-L
1. Adds Complexity/Failure Modes	2nd valve increases both likelihood and rate of depressurization.	2nd O-ring forces designers to extend SRB tang, reducing primary and secondary O-ring effectiveness.
2. Increases Opacity	Cosmonauts did not know which valve to close during depressurization of cabin.	Managers mistakenly believed the SRB joint was redundant.
3. Encourages Risk	Cosmonauts did not wear pressure suits during reentry.	Challenger was launched despite record cold temperatures.
4. Discourages further optimization of single element hardware	N/A	SRB flew despite evidence of past O-ring erosion

Table 1. Limitations of redundancy as exemplified by Soyuz 11 and Challenger STS-51-L.

IV. Conclusions

The previous case studies highlight instances where redundant elements have contributed to Loss of Crew, either by failing catastrophically (Soyuz 11) or by failing to perform as designed (*Challenger* STS-51-L).

These findings should not be taken as evidence that redundant elements are inherently unsafe. A secondary O-ring may not have prevented *Challenger*, but it prevented a *Challenger*-like accident on STS-51-B[13] (when a primary O-ring failed to seal). In a similar vein, a backup pressure relief valve may have condemned the crew of Soyuz 11, but it saved the cosmonauts on board Soyuz 23, providing fresh air to the capsule after its landing and partial submersion in Lake Tengiz[28].

Instead, the findings presented here suggest that redundancy and its correlation with safety—small or large, direct or inverse—is based on a network of variables whose interactions are difficult to identify and virtually impossible to model. Therefore, design requirements that mandate set levels of redundancy—though written to reduce risk—may not always do so.

Acknowledgments

The author would like to acknowledge Dr. David Klaus for reviewing a draft of this article.

References

[1]FAA, "FAA Systems Safety Handbook," 2000.
[2]NASA, "NASA General Safety Program Requirements," NPR 8715.3C, 2008.
[3]FAA, "Established Practices for Human Space Flight Occupant Safety-Draft," 2013.
[4]NASA, "International Space Station (ISS) to Commercial Orbital Transportation Services (COTS) Interface Requirements Document (IRD)," SSP 50808, 2012.
[5]Butler, R. W. "A primer on architectural level fault tolerance," NASA TM-2008-215108, 2008.
[6]Downer, J. *When failure is an option: Redundancy, reliability and regulation in complex technical systems,* London School of Economics and Political Science, London, 2009.
[7]Low, G. M., "What Made Apollo a Success?" *Astronautics and Aeronautics*, Vol. 8, 1970, pp. 36-45.
[8]Hitt, D., Garriott, O. & Kerwin, J. *Homesteading Space: The Skylab Story*, University of Nebraska Press, Lincoln, NE, 2008.
[9]Ocampo, R. P., "PhD Dissertation", Aerospace Engineering Sciences, University of Colorado, Boulder, CO (to be published).
[10]NASA. "Report of Apollo 13 Review Board," NASA-TMX-65270, 1970.
[11]Ivanovich, G. S. *Salyut-The First Space Station: Triumph and Tragedy*, Praxis, Berlin, 2008.
[12]Kamanin, N., "Diary of Nikolai Petrovich Kamanin", 1971.
[13]Rogers et al., "Report to the President by the Presidential Commission on the Space Shuttle Challenger Accident, 1986.
[14]U.S. House of Representatives. "Investigation of the Challenger Accident," House Report 99-1016, 1986.
[15]NASA, "SRM Clevis Joint Leakage Study," 1977.
[16]Thiokol, "Analytical Evaluation of the Space Shuttle Solid Rocket Motor Tang/Clevis Joint Behavior," PC 102302, 1978.
[17]Greenfield, M. A., "Normal Accident Theory: The Changing Face of NASA and Aerospace," Presentation, 1998.
[18]Perrow, C. *Normal accidents: Living with high risk technologies*. Basic Books, New York, 1984.
[19]Sagan, S. D., *The limits of safety*, Vol. 9, Princeton University Press, Princeton, NJ, 1993.

[20]Leveson, N., Dulac, N., Marais, K. & Carroll, J., "Moving beyond normal accidents and high reliability organizations: a systems approach to safety in complex systems," *Organization Studies*, Vol. 30, 2009, pp. 227-249.
[21]Stevens, W. K., "Man has yet to Master Shuttle's Sophistication," *The New York Times*, November 15, 1981.
[22]Haskins, C., Forsberg, K., Krueger, M. Walden, D., & Hamelin, D., *Systems Engineering Handbook*, INCOSE, Seattle, 2011.
[23]Thiokol, "Presidential Commission Development and Production Panel, Response to Panel Question/Special Actions- SRM and Titan III Clevis Joint Comparison," PC 073979, 1986.
[24]NASA, "SRB Critical Items List,", 1982.
[25]Vaughan, D., *The Challenger launch decision: Risky technology, culture, and deviance at NASA*, University of Chicago Press, Chicago, 1997.
[26]Sagan, S. D., "The Problem of Redundancy Problem: Why More Nuclear Security Forces May Produce Less Nuclear Security," *Risk Analysis* Vol. 24, 2004, pp. 935-946.
[27]NASA "STS 51-E Flight Readiness Review, Level 1," 1985.
[28]Portree, D. S. F., "Mir hardware heritage," NASA RP 1357, 1995.