

THE CONSTRUCTAL ANALYSIS OF WARFARE

G. WEINERTH

Duke University, Pratt School of Engineering, USA.

ABSTRACT

Military science has for millennia tried to study and maximize a fighting force's advantages in battle. However, it has never employed mathematical tools in its analysis. Constructal theory provides a means to understand the optimal placement, structure, and deployment of phalanx based armies. Because phalanxes possess a direct geometric correlation between strength and vulnerability, their form is well suited to theoretical analysis. The optimum phalanx formation is given as one that minimizes exposed flanks while maximizing the unit's effective front and deployment depth. In a further example, a given hypothetical distribution of fast, slow, and medium troops charts the development and evolution of flanking maneuvers and ultimately Hannibal's double envelopment strategy at Cannae.

Keywords: constructal theory, envelopment, hoplite, phalanx, strategy, tactics, troop deployment, war.

1 INTRODUCTION

Victorious warriors win first and then go to war, while defeated warriors go to war first and then seek to win.

Sun Tzu.

Military science has never viewed war as a mathematically optimized flow system. But that is exactly what warfare is. All combatants seek to improve upon their own strengths at certain positions while exploiting their enemy's weaknesses in a constant state of flux. This paper attempts to optimize strategy and tactics in ancient warfare through the use of Bejan's Constructal Theory.

2 BASIC PRINCIPLES

The two major doctrines involved are the principles of encirclement and concentration of force. However, they each have a unique geometrical and physical representation [1].

2.1 Encirclement

The tactic of encirclement is one of the oldest and most basic to warfare (Fig. 1). It is evident in pincer movements, double envelopments, and flanking maneuvers. Each is a means of attacking an enemy on their sides or rear while still being frontally engaged. No fighting force can bring their attention to two fronts without dividing their strength. Encirclement also has the advantage of limiting the enemy's escape routes, lines of communication, and logistical support. The weakness of encirclement is that the tactic requires a thinner and weaker spread of troops necessary to surround the enemy, which could be susceptible to a countermeasure of a focused breakout.

2.2 Concentration of force

The second tactical principle is that of concentration of force (Fig. 2). Essentially, this is a means of using a greater force against a weaker force – to 'divide and conquer'. The best possible offensive move is to bring all forces to fight against a limited fraction of the enemies' number. For example, if I have a total of 100 soldiers, and my opponent has 300, I should arrange my plan of battle so that the full force of my 100 comes against successive units of 50 enemy soldiers. By concentrating my force against the divided enemy, I increase the possibility of winning those smaller engagements.

This principle features heavily where geography hinders one side's freedom. Mountain passes are easily secured by a defensive force since the terrain requires a larger enemy to occupy a limiting area (Fig. 3). In this manner, a minimal Spartan force was able to hold off a vastly superior Persian army at the Battle of Thermopylae Pass. Similarly, Sun Tzu writes that when an invading force crosses a river, a commander should let half the opposing army cross before attacking so that the enemy's force is divided and pinned by the river [2].

3 EXAMPLES FROM HISTORY

The Greek phalanx provides an excellent starting point for historical analysis. The phalanx was composed of *hoplites*, slow moving heavy infantry. Phalanx engagements were essentially pushing matches. The less powerful side would eventually break ranks and flee. The phalanx's front was critical. The side with the shorter front risked being overlapped at the ends and enveloped by an enemy with a longer front. However, a commander could not thin his ranks too much. A phalanx's depth was directly correlated with its pushing power and resilience. A narrow phalanx would easily break when faced with an enemy phalanx possessing greater depth and pushing power.

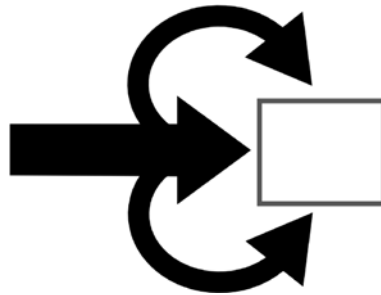


Figure 1: Basic encirclement maneuver.

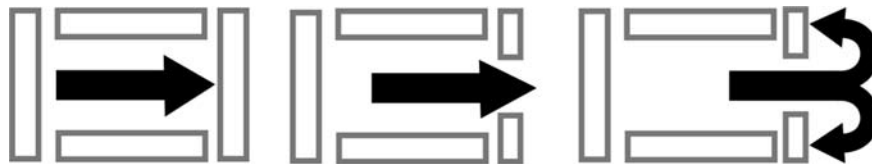


Figure 2: Breakout from encirclement through a concentration of force.

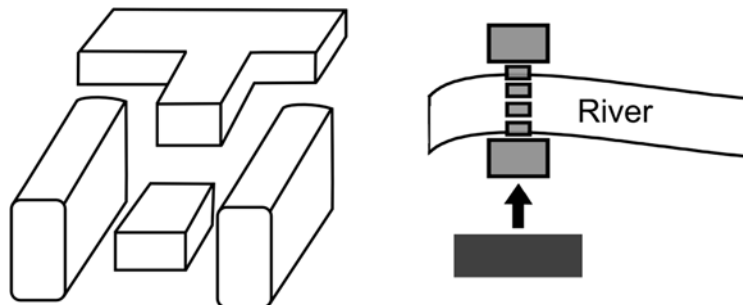


Figure 3: A mountain pass limits numerical superiority (left). Terrain advantage in crossing a river (right).

In addition, the phalanx had two principal shortcomings. One was its limited mobility, and the other was its vulnerability on its flanks and rear. Despite these drawbacks, the phalanx form did not impede tactical brilliance or innovation.

3.1 The Battle of Marathon – 490 BC

The first major tactical innovation in phalanx warfare was evidenced at the famed Battle of Marathon. The Persian army vastly outnumbered the Athenians and included archers and cavalry troops, divisions which Athenians lacked.

The Greek commander, Miltiades, realized that the front line of the Persians was longer than that which the Athenians could successfully match. In order to create an effective front equal to that of the Persians, Miltiades thinned his center lines and concentrated his own flanks.

3.1.1 Strategy at Marathon

The Athenians seized the initiative by charging at the Persian line. However, the Persians countered and advanced upon the apparent weakness of Greek thin center. This caused the Greek line to fall inward. The flanks, however, stood strong. Then, when the Persian infantry had pushed the weak Greek center back, the Athenian commander rallied and stopped the Persian advance. The Greek flanks pivoted inwards and intercepted the Persian line, enveloping them on three fronts ([3], p. 133).

This strategy artfully changed the reinforced flanks from a liability to a tactical advantage. By merely turning to one side, the narrower Athenian flanks became wider attacking wings on the weak Persian army's flanks (Fig. 4).

3.2 Envelopment at the Battle of Cannae – 215 BC

Hannibal's famous march through the Alps was a bold invasion of the Italian Peninsula, allowing him to challenge Rome directly. The Roman general Varro intended to destroy Hannibal with an army that more than doubled the Carthaginian force. However, Hannibal's superior cavalry and the diversity of his army proved invaluable to his tactical success.

Varro arranged his army in a standard position of three lines. Hannibal also arranged his army in a row, but with a specific composition. The center was made up of his weakest light infantry, directly facing the Roman line. On the extremes of his infantry line, he placed his stronger phalanxes. Finally, the cavalry set beyond the line on either side of the infantry.

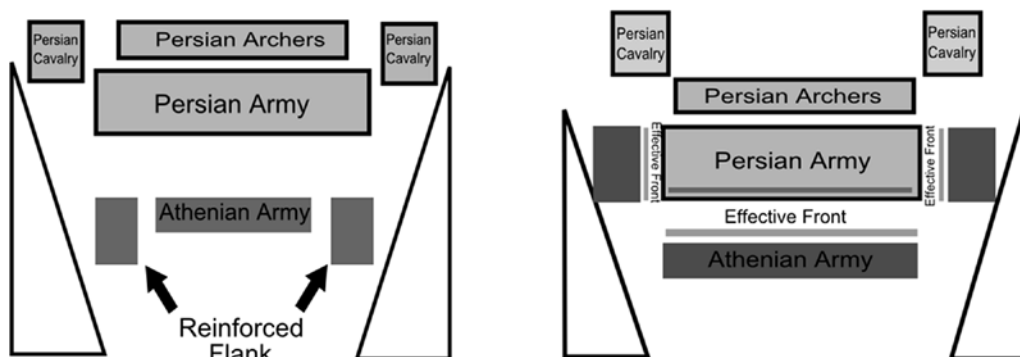


Figure 4: Battle of Marathon – opening positions (left) Athenian envelopment (right).

3.2.1 Strategy at Cannae

At the beginning of the battle, the Carthaginian cavalry quickly drove off the Roman cavalry. The Roman infantry, however, found better luck against the weak center of the Carthaginian line and it began to be pushed backwards. Despite this, the stronger infantry anchoring the ends of the Carthaginian line held firm. Hannibal intended the center to experience a controlled withdrawal as the Romans drew deeper into the Carthaginian line. When they had advanced to a critical point, Hannibal's heavy infantry turned 90° and were in immediate position to face the vulnerable Roman flanks. The Roman infantry was suddenly hemmed in on three sides. The weak center was then rallied by Hannibal himself and managed to hold firm as the rest of the Carthaginian line closed in on the Roman flanks (Fig. 5).

Once the Romans were enclosed on three sides, the Carthaginian cavalry came in on the infantry's rear, entirely surrounding them. The Roman army was massacred where it stood ([3], pp. 143–145, 307).

3.3 Concentration of force in phalanx warfare

At the Battle of Leuctra (371 BC), the city state of Thebes was poised to counter the military dominance of Sparta with an unprecedented phalanx formation. The Theban front and center were thinned and extended in order to heavily reinforce the left flank. The reorganized left flank was double the strength of the center and right combined. The right and center lines themselves were arranged in a staggered echelon formation. The Theban army was the first to purposefully create a concentration of force in an oblique deployment (Fig. 6). In battle, the reinforced left flank

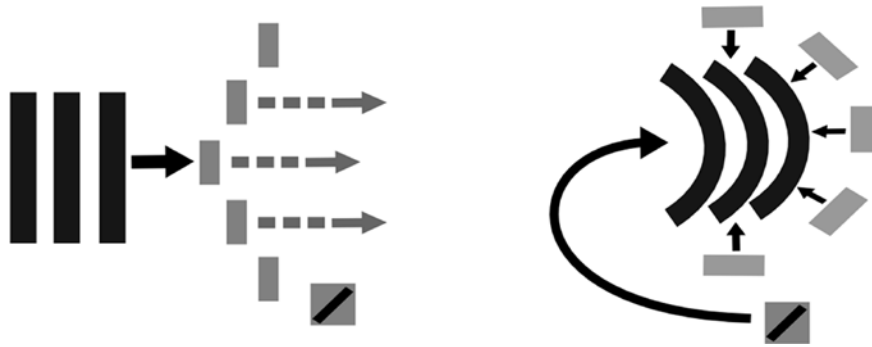


Figure 5: The calculated withdrawal of the Carthaginian line (left). Hannibal's rally and double envelopment (right).

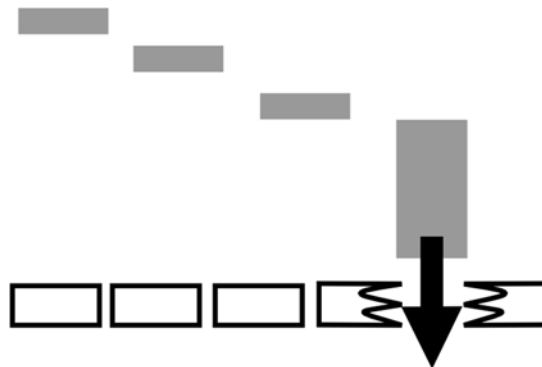


Figure 6: Theban concentration of force.

overpowered the Spartan line with its tremendous momentum and defeated the most accomplished hoplite army of the ancient world [4].

3.4 Concentration of force in Roman infantry maneuvers

The Roman legionnaire was drilled to change positions and attack with an unprecedented synchronization. This mobility allowed for a new degree of creativity in Roman concentration of force. To crash through their enemy's lines, they could assemble in a wedge shape called a *cuneus* which applied maximum pressure to penetrate the enemy's ranks (Fig. 7). Another Roman tactic involved 'rolling up the ranks'. Like the Battle of Leuctra, a reinforced oblique column would be sent to break through the enemy's lines. This column, once breaking through, would then move to attack the rear, producing a two front battle. This two-phase attack perfectly incorporates both envelopment and concentration of force principles ([3], pp. 332, 417).

4 CONSTRUCTAL ANALYSIS

Constructal theory states that shapes and specific geometry arise from a flow system evolving in time subject to specific constraints [5]. Subject to geographical and physical constraints, the armed force is a system which seeks to flow to vulnerable parts of the enemy formation and avoid stronger areas of the enemy's position.

The correlation between area and army size is a simple matter of recognizing that there are only a set amount of soldiers available to each side. Because phalanx formations were based on regular geometric spacing, it is both convenient and accurate to diagram armies as varying, moving areas. Each unit has a fixed speed in relation to the division type. Heavy infantry was usually the slowest, light infantry moderately faster, while cavalry was the fastest. Because of these varying speeds, they each played a very different role in combat.

4.1 Basic geometry and posture in the Greek phalanx

Relative 'strength' and 'weakness' of a phalanx unit was directly correlated to its shape. The deeper the phalanx or the more cavalry present in a contingent gives a primary indication of how strong it was. The phalanx had to find a balance between depth (strength) and width (effective fighting front). The side with the shorter effective fighting front risked being overlapped and enveloped by the enemy. Yet, a line that was too long sacrificed frontal pushing power. With a fixed number of troops, this presented an inherent tradeoff between power and vulnerability.

Much like the constructal analysis of river channels, phalanx geometry can be optimized to have the best effective front/vulnerable front ratio while maintaining strength (depth) (Fig. 8).

Similar to the analysis applied to the cross-sections of open channels, there are optimal shapes for a specified area with given limitations. Defensive and aggressive postures require different ratios of

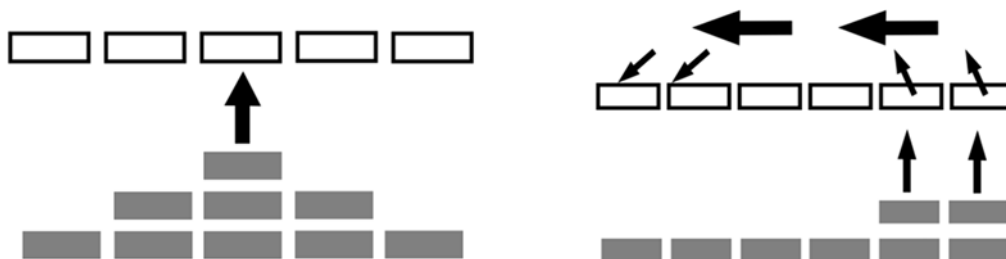


Figure 7: Roman *Cuneus* formation (left) 'rolling up the ranks' (right).

effective and vulnerable fronts to total area. Table 1 provides a comparison for differing phalanx formations. The optimal formation would use the maximum offensive effectiveness of the wedge for offense and the defensive properties of the semi-circle. This outcome reflects both the practice and theory of optimal distribution of vulnerabilities and strengths (Fig. 9). Also noted is the reappearance of the semi-circle as possessing the most efficient geometry as a ‘constructal’ shape [6].

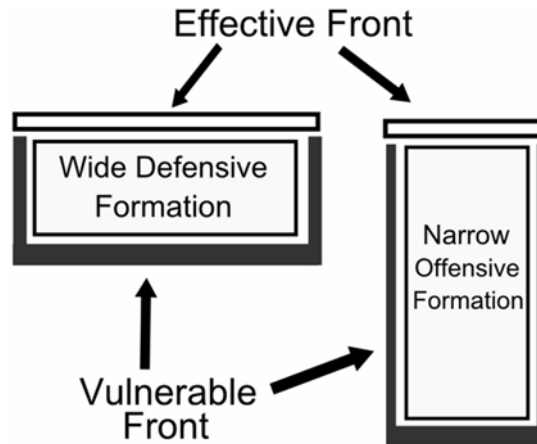


Figure 8: Area and geometrical considerations.

Table 1: Comparison of phalanx shapes.

Shape	Effective front	Depth (strength)	Vulnerable front	Effective front to vulnerable front ratio	Strength to vulnerable front ratio
1. Rectangle	W	D	$W + 2D$	$\frac{D}{2D + W}$	$\frac{D}{2D + W}$
2. V-shape	W	D	$2\sqrt{D^2 + \frac{W^2}{4}}$	$\frac{W}{2\sqrt{D^2 + \frac{W^2}{4}}}$	$\frac{D}{2\sqrt{D^2 + \frac{W^2}{4}}}$
3. Wedge	$2\sqrt{D^2 + \frac{W^2}{4}}$	D	W	$\frac{W}{2\sqrt{D^2 + \frac{W^2}{4}}}$	$\frac{D}{W}$
4. Semi-circle	D	$\frac{D}{2}$	$\frac{\pi D}{2}$	$\frac{2}{\pi} = 0.6366$	$\frac{1}{\pi} = 0.318$

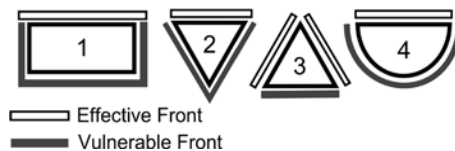


Figure 9: Corresponding diagram for shapes in Table 1.

4.2 Constructal derivation of Marathon and Cannae envelopments

First, consider the best way for an army, comprised of slow, medium, and fast troops ($V_0 < V_1 < V_2$) with unknown distances W_1 , W_2 and W_3 between them to arrange in a line before encircling a stationary enemy camp. The enemy camp has a set length and width of L_E and W_E and is at a distance D from the line (Fig. 10). Each unit traverses the shortest distance possible to reach the enemy camp. It is expected that the V_1 and V_2 units will take diagonal routes to the sides and rear of the enemy camp and the V_0 unit will travel in a straight line to the front of the enemy camp. The equations model these diagonal paths as the hypotenuses of right triangles with base measurements of W distances. However, the W_2 and W_3 positions are to be determined so that all of the divisions will encircle the enemy at the *same time* so as to coordinate the attack and maximize the element of surprise. Setting time as constant will ensure that the slowest division (V_0) will arrive at the enemy camp front (D) at the same time the fastest (V_3) will appear at the enemy camp rear ($D + L_E$). The following equations can therefore be arrived at:

$$\text{Time} = \frac{D}{V_0} = \frac{\sqrt{\left(D + \left(\frac{L_E}{2}\right)\right)^2 + W_2^2}}{V_1} = \frac{\sqrt{(D + L_E)^2 + W_3^2}}{V_2} \quad (1)$$

With optimal spacing of W_2 and W_3 :

$$W_2 = \sqrt{\left(\frac{V_1 D}{V_0}\right)^2 - \left(D + \frac{L_E}{2}\right)^2}$$

$$W_3 = \sqrt{\left(\frac{V_2 D}{V_0}\right)^2 - (D + L_E)^2} \quad (2)$$

This, as expected, gives a spacing value dependent on troop speed and the enemy camp's size and distance from the front line.

Suppose that instead of being stationary, the enemy is a phalanx moving with speed V_E (Fig. 11). The principle distance will vary as the two armies approach each other, $D = (V_0 - V_E)t$. This distance is dependent on the relative speeds of the enemy and the slowest company. Inserting this value into

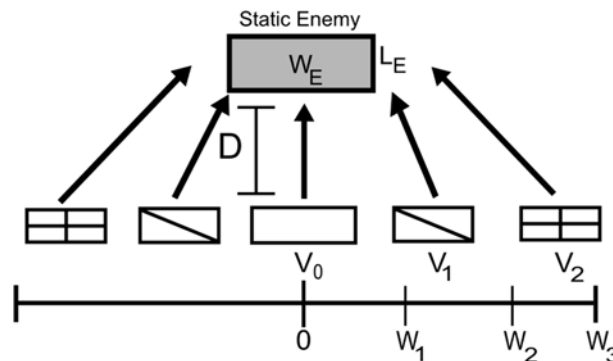


Figure 10: Straight line paths to surround a static enemy.

the previous equations for W_2 and W_3 and differentiating with respect to time for the minimal distance traveled for each group gives a very interesting solution. The shortest time is when the slowest company *retreats*, drawing in the enemy towards a gap in the line, where it will be horizontal to the V_1 company.

At a certain time scaled to

$$\frac{1}{2} \frac{V_0^2 L_E}{V_1^2 V_0 + V_1^2 V_E + V_0^3 - V_0^2 V_E},$$

the V_1 force, positioned at

$$W_1 = \frac{1}{2} \frac{V_1 V_0^2 L_E}{V_1^2 V_0 + V_1^2 V_E + V_0^3 - V_0^2 V_E},$$

would be at optimum position to attack on the sides. A similar line of reasoning will give results for a reduced and maximized position for the V_2 company to attack on the rear of the advancing enemy. Therefore, it can be seen that the brilliant double envelopments at Marathon and Cannae can be obtained simply by optimizing the positions of each unit with respect to which fronts they are best suited to engage (Fig. 12).

5 CONCLUSION

This paper attempted to examine how to optimize the abstract system of ancient warfare on mathematical principles. The solutions are rather surprising and confirm much that has been predicted by

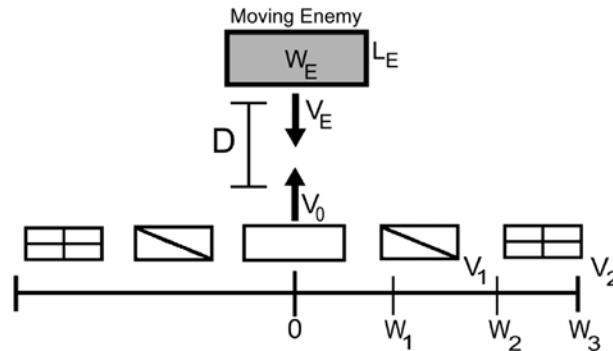


Figure 11: Distributed army facing a moving enemy.

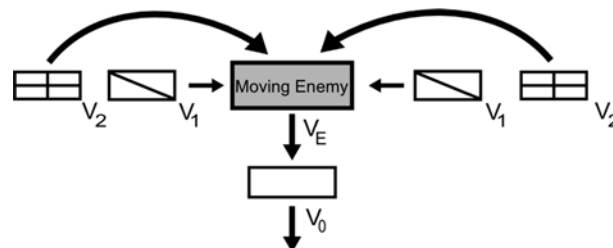


Figure 12: Double envelopment achieved by optimization.

constructal theory as to optimal geometry and distributions. Efficient design is brought about through the efficient distribution of imperfections within an evolving system. A fighting force, like any other system, seeks to minimize its inherent weaknesses and exploit the vulnerabilities of the enemy. Future study may provide improved analyses, but the goal should not be to perfect war. History abounds with conflicts wherein superior strategy does not result in ultimate victory. However, it is hoped that this framework gives a more rational understanding to a widespread and important social phenomena. It is only appropriate for science and mathematics to provide insight into the nature of warfare. Once knowledge of battles and engagements are reduced to a science rather than elevated as an art, there can be further progress towards the elimination of war.

REFERENCES

- [1] Bellamy, C., *The Evolution of Modern Land Warfare: Theory and Practice*, Routledge: NY, pp. 17–23, 1990.
- [2] Sun, T., *The Art of War*, (trans. Lionel Giles), Barnes and Nobles Classics: NY, p. 43, 1993.
- [3] Gabriel, R.A., & Boose, D.W., *The Great Battles of Antiquity*, Greenwood Press: Westport Connecticut, 1994.
- [4] Bradford, A.S., *With Arrow, Sword, and Spear. A History of Warfare in the Ancient World*, Praeger Press: Westport Connecticut, pp. 89–99, 2001.
- [5] Bejan, A., *Shape and Structure, from Engineering to Nature*, Cambridge University Press: Cambridge, UK, p. 324, 2000.
- [6] Bejan, A. & Lorente, S., *Design with Constructal Theory*. Draft, Chapter 3 Section 2, 2007.