Coupled Analysis of Heat Transfer and Temperature Dependent Plastic Deformation in Friction Stir Welding

1,a Patricio F. Mendez and 2,b Karem E. Tello

1 University of Alberta, Edmonton, AB, T6G 2V4, Canada
2 Colorado School of Mines, Golden, CO, 80401, USA
a pmendez@ualberta.ca, b ktello@mines.edu

Abstract
This paper aims at summarizing previous work on fundamental estimations about the FSW process with a minimum of ad-hoc experimental input. To do so, an approach based on asymptotic analysis, along the lines of boundary layer analysis in fluid mechanics is developed, resulting in closed-form expressions of great generality. The scaling laws obtained predict the correct trends and orders of magnitude for the maximum temperature reached in the process, the thickness of the shear layer, the shear stress around the pin, the torque and the thermal effect of the shoulder. A comparison of the scaling laws obtained with published experimental measurements and numerical simulations show that within the range of validity, the expressions obtained captures the right order of magnitude and trends with no model calibration. In addition to process predictions, the analysis also yields dimensionless groups, which can be used to relate existing experiments to theoretical predictions, thus calibrating the models in an accurate and rigorous way. The available experimental and numerical data were used to calibrate the closed-form predictions resulting in simple and accurate closed-form expressions (14% standard deviation from measurements for maximum temperature). The proposed calibrated expressions are accurate for a broad range of materials including aluminum alloys 6061, 2024, 7075, 7050, 5083; mild steel, stainless steel, and titanium. This model aims to help metallurgical research on FSW by providing simple and accurate approximations of broad validity across alloy systems that can be used to quickly assess problem parameters for models of phase transformations, to generalize observations about defect formation, and to design FSW applications of radically different materials or dimensions than currently known.

1. Introduction

1.1 Friction Stir Welding
Friction Stir Welding (FSW) is a joining process invented at The Welding Institute (TWI) in 1991 [1], and is considered a solid state joining process where no melting is expected at any point of the base plate. Figure 1 presents a schematic of FSW. In this process, a tool rotating with its axis near perpendicular to the base plates penetrates and deforms them in such a way that independent pieces of metal become locally mixed and joined. The tool consists of a shoulder and a pin, in which the shoulder typically has a diameter three times the diameter of the pin, and the length of the pin is slightly shorter than the thickness of the plate. The base plates are often welded in the butt joint configuration.

Figure 1: Schematic of Friction Stir Welding[2]

The heat generated increases the temperature in the base plate usually below its melting temperature, softens it, and
facilitates the flow of material around the tool to consolidate the joint. Although FSW was first used to join aluminum and its alloys, currently it is possible to join mild steels, stainless steel, magnesium alloys, titanium alloys among others materials. Some applications of FSW are Delta II and Delta IV expendable launch vehicles from Boeing and the external tank for the Space Shuttle made of AA2195 aluminum alloy [3].

1.2 Scaling Analysis

FSW is a very complex coupled problem. What makes the coupling especially complex in this case is the variation of mechanical properties with temperature. Temperature dependent properties often complicate modeling efforts, and for the case of FSW, the softening with temperature, which results in a localized shear region, is the key concept enabling the process.

FSW, like all other engineering process can be described by a set of partial differential governing equations; alas, the proper formulation of the problem is near intractable. Also, even if an accurate solution was found numerically or experimentally, the validity of this solution would be only for the parameters considered. Any variation of parameters would require a sensitivity study. Sensitivity studies in problems with poorly understood parameters (or with too many parameters) can be misleading, and must be supplemented with additional insight. This paper aims at contributing with the additional insight that will enhance numerical models, experiments, and help guide sensitivity studies.

In the search for insight, we aim at developing a FSW model "as simple as possible, but not simpler"[4]. This will mean in this case, to obtain power laws through the technique of asymptotic scaling studied in depth in[5].

The technique of scaling is especially suited for FSW and other engineering problems described by a system of partial differential equations, which present some challenges such as i) uncertainty, about their validity, ii) difficulty in solving, and iii) the generalization of the results is difficult even when they are solved. Scaling solutions have the following characteristics: i) they are achieved with little computational time and complexity ii) they are expressed in closed-form iii) their range of validity is clearly determined iv) they discriminate the dominant factor from the less relevant ones. The utility of a scaling approach for welding research has recently been addressed in[6].

2. Modeling Considerations

To make the problem tractable, the scope of the model presented here is limited to the deformed area surrounding the pin. It will be shown later that for many practical cases, the conclusions obtained apply to the whole process, shoulder included.

The approach considered here is that first introduced in[7], and it is based on Prandtl’s approach to the boundary layer problem in fluid mechanics[8, 9]. The approach consists of realizing that the continuum of material surrounding the pin can be divided into two distinct regions separated by a diffuse (but narrow) transition. In this case, the “inner region” near the pin will receive a different mathematical treatment than the “outer region” far from the pin.

In the inner region, the pin is surrounded by the “shear layer” (also called the “mixed zone” and the “deformation zone.” Within the shear layer, significant plastic deformation is observed, while outside the shear layer (the “outer region,” the substrate remains essentially undeformed. The presence of a thin shear layer surrounding the pin is analogous to the viscous boundary layer surrounding a body moving through a fluid. In fluid mechanics, when inertial forces are dominant in a fluid around a body, viscous forces are still dominant in a thin region immediately close to the body; this region is the viscous boundary layer. The concept of boundary layer enables insightful, efficient and fast solutions of fluid mechanics problems that would otherwise be very challenging, for this reason Prandtl’s original paper is considered one of the most important in the history of fluid mechanics[10].

In FSW the thin shear layer is analogous to the viscous boundary layer, and can be defined as the region in which heat generation from plastic deformation is dominant. The region of the substrate
outside the shear layer is dominated by heat conduction; this region is the analogue of the external flow in the viscous boundary. In this approach, a different asymptotic solution is found for each region, and both solutions must be matched at their boundary.

For generality of problem statement, a cylindrical, threadless pin is considered. Acknowledging that different pin geometries can have a significant influence on the resulting weld, the analysis performed still captures the fundamental mechanisms of coupled heat and plastic deformation shared by all FSW processes.

FSW can be performed with a varying degree of success in a variety of regimes, where different physical phenomena have widely varying degrees of influence. The analysis followed here considers four conditions that are most typical in successful practical implementations. These conditions were tested to be valid with the use of self-consistent scaling considerations too:

(i) the travelling pin can be considered a steady state, slow moving heat source
(ii) there is small advance per revolution
(iii) the shear layer is thin
(iv) the heat from the shoulder is not the dominant factor determining the peak temperature.

3. Scaled Governing Equations

3.1 Heat conduction in the shear layer

The consideration that the shear layer is thin results, just like in fluid mechanics, in a nearly one-dimensional transport problem (in Prandtl’s analysis, variations in the longitudinal direction are negligible with those transverse to the boundary layer). This results in the following heat transfer equation:

\[ \frac{d^2 T}{dx^2} + \frac{q}{k} = 0 \]

Eq. 1

Which can be scaled using the concepts from[5] as

\[ -2 \frac{\Delta T_s}{\delta^2} + \frac{\dot{q}_c}{k_0} = 0 \]

Eq. 2

where \( \delta \) is the thickness of the shear layer, and the notation using a hat indicates estimated characteristic values, and \( q \) is the volumetric heat generation and \( k \) the thermal conductivity of the base metal.

3.2 Heat generation in the shear layer

The heat generation in the shear layer can be captured as

\[ q(x) = -\eta x \frac{d\dot{v}}{dx} \]

Eq. 3

where \( \eta \) is an efficiency, \( \tau_c \) is a characteristic shear stress, \( v \) is velocity of the material being deformed, and \( x \) is the coordinate normal to the shear layer (i.e. radial coordinate with the origin at the pin/substrate interface). Scaling of this equation results in

\[ \dot{q}_c = \frac{3}{2} \eta \frac{\omega a}{\delta} \]

Eq. 4

where the differential expression is estimated based on the rotation velocity, pin radius, and shear layer thickness.

3.3 Constitutive equation in the shear layer

The Sellars and Teggart[11] constitutive model in the high temperature, high strain rate is applicable to this problem, and has the following expression:

\[ \dot{\varepsilon} = A \left( \frac{\tau}{\tau_R} \right)^n \exp \left( -\frac{Q}{RT} \right) \]

Eq. 5

which can be scaled into

\[ \frac{3}{2} \frac{\omega a}{\delta} \approx AB \left( \frac{\tau_c}{\tau_R} \right)^n \frac{\Delta T_s}{\Delta T_m} \]

Eq. 6

where

\[ \Delta T_s = T_s - T_0 \]

\[ \Delta T_m = T_m - T_0 \]

\[ B = \exp \left( -\frac{Q}{RT_m} \right) \]

all notation is described in detail in[7], however, it is worth highlighting that the scaled equation interprets the Arrhenius-type evolution as a steep linear variation of
shear stress with temperature near the melting point.

The value of the parameters for the constitutive model were obtained from raw experimental data for high-strain rate, high temperature deformation reported in the literature and summarized in [12]. For example, the fitted constitutive model for aluminum 6061 is represented in Figure 2. In this figure the low temperature regime is also considered, resulting in a “sinh” function being part of the model used for fitting.

![Figure 2: Fitted constitutive model and constants for aluminum 6061[12]. The solid dots represents measurements reported in[13]. The dashed lines are older model constants reported in[14]](image)

3.4 Heat conduction outside the shear layer

This is the “outer region,” where heat is transmitted by conduction with little heat generation. Decoupling the FSW model into inner and outer regions enabled the consideration of different thermal boundary conditions on the substrate, such as convective cooling to the atmosphere, and conduction cooling to the backing plate.

The hypothesis of a slow-moving heat source implies that the isotherm delimiting the shear layer is approximately circular and concentric with the pin. In these conditions, Rosenthal’s thin plate solution[15] is applicable:

\[
T(x') - T_p = \frac{q_f}{2\pi k_0} \exp \left( -\frac{Vx'}{2z_0} \right) K_0(\xi) \tag{Eq. 7}
\]

all notation is explained in detail in [7], but it is interesting to highlight that this equation is capable of capturing the heat loses to the backing plate when that is necessary. Scaling of this equation considering the torque generated by the pin results in:

\[
\Delta T_0 = T_0 - T_\infty = \eta \frac{\omega a^2}{k_0} \hat{T}_0(Pe) \tag{Eq. 8}
\]

4. Scaling Expressions

Equations 2, 4, 6, 8 involve many parameters known a-priori, but only four unknown estimations: the characteristic shear stress \( \hat{\tau}_c \), the shear layer thickness \( \hat{\delta} \), the maximum temperature \( \hat{T}_m \), and the volumetric heat generation \( \hat{q}_c \). The solution of this system of four equations with four unknowns result in the following expressions[6, 7]:

\[
\hat{T}_m = T_0 + \Delta T_m \left[ \frac{3}{2AB\Delta T_m} \left( \frac{\eta K_0}{\Delta T_0} \right) \left( \frac{a^2 \tau_c}{k_0} \right)^{n+1} \left( \frac{\omega}{\eta K_0} \right)^{n+1} \right]^{\frac{1}{2}} \tag{Eq. 9}
\]

\[
\hat{\tau}_c = \frac{k_0 \Delta T_0}{\eta \omega a^2 K_0} \tag{Eq. 10}
\]

\[
\hat{\delta} = \frac{8}{3} \Delta T_m \left( \frac{\eta K_0}{\Delta T_0} \right)^{n+1} \left( \frac{\omega}{\eta K_0} \right)^{n+1} \tag{Eq. 11}
\]

\[
\hat{q}_c = \frac{3}{4} \frac{3}{\Delta T_m} \left( \frac{\eta K_0}{\Delta T_0} \right)^{n+1} \left( \frac{\omega}{\eta K_0} \right)^{n+1} \tag{Eq. 12}
\]

From these expressions many other useful ones can be derived, for example shear rate

\[
\hat{\gamma} = 2 \frac{\omega a}{\hat{\delta}} \tag{Eq. 13}
\]

or heat input

\[
\hat{H}.\Phi. = \frac{M \omega}{V} \tag{Eq. 14}
\]

For both cases of Eq. 13 and 14, what is remarkable is that they can be estimated before any experiment is performed, with no need to measure torque or temperature evolution.
5. Validation of Scaling Laws

The simplicity of equations 9-12 is paradoxical, when considering the complexity of FSW, it is fair to ask oneself, are these expressions valid? In what range? These expressions were validated against all available published results at the time and presented in [16]. The validation of the expression for maximum temperature is the one that best illustrates the significance of the accomplishments. For the comparison, the parameter \( f_T \) will be defined

\[
\frac{T_T - T_\infty}{T_T - T_s} = f_T
\]

This parameter relates the measured temperature to the estimation based on Eq. 9. Appropriate scaling should result in values of \( f_T \) of approximately 1, and indeed that is the case.

Figures 3-6 compare the scaling prediction of maximum temperature (Eq. 9) with published experimental and numerical values. The vertical axis in each figure represents the parameter \( f_T \). The horizontal axes are dimensionless groups representing the relative magnitude of the physical phenomena neglected in the four approximations restricting the scope of the model. The values on the horizontal axis can be calculated from problem parameters without the need for experiments. A value of one on the horizontal axis means that the neglected effect is of the same order as the effect considered dominant.

In all cases the estimated temperatures are of the right order of magnitude (\( f_T \sim 1 \)) in the range of validity of the hypotheses. It is remarkable that this result was obtained with no correction factors and no FSW specific empirical inputs.

There are clear limits to the validity of the approximations made, indicated by a value of one on the horizontal axis, but surprisingly, the scaling law still predicts the behaviour of the system beyond the limit of validity for approximations I, II, and IV. This is a welcome result that could not be anticipated by the scaling theory alone, and emphasizes the robustness of the scaling approach.

5.1 Sources of error

The errors in the scaling predictions have four sources[6]: error in the published values, approximations in the physics, approximations in the mathematical treatment, and errors in the materials properties used for the scaling predictions.

The error in the published values are inherent to the nature of measurements and simulations and are expected to be small. These errors result in scatter and are interpreted as “random” in their modeling. Errors from approximations in the physics originates in physical phenomena that are known to exist but are neglected. When scaling is performed properly, error from simplified physics are also frequently considered random. In contrast with the previous two error sources, simplifications in the mathematics cause systematic error. Error in the materials properties used as input show as alloy dependent systematic error.

![Figure 3: Dimensionless maximum temperature in base plate versus approximation (i) captured as a Peclet number[7]](image-url)
Figure 4: Dimensionless maximum temperature in base plate versus approximation (ii) captured as the ratio between translation and rotation speed in tool[7]

Figure 5: Dimensionless maximum temperature in base plate versus approximation (iii) captured as the dimensionless thickness of shear layer[7]

Figure 6: Dimensionless maximum temperature in base plate versus approximation (iv) captured as the ratio of temperature increase due to shoulder relative to peak temperature increase[7]

6. Calibration of Scaling Laws

In the scaling analysis presented, the scaling laws were presented in the context of hypothesis of approximation. Dimensionless groups resulting from the scaling analysis quantify the degree of fulfillment of the required hypotheses. These dimensionless groups are the basis of rigorous calibration of the scaling laws. The comparison between scaling predictions and measurements result in a much nuanced activity than the frequently used (yet despised when used by others) “fudge factor. In this work, there is no constant calibration factor, but a function that takes into account the degree of fulfillment of the relevant hypotheses. This analysis was performed for all scaling estimations in [16], but will be exemplified here only for the temperature estimates.

Inspection of Figs. 3–6 shows that a correction factor should depend strongly on the degree of fulfillment of hypothesis iii, expressed as the ratio between the estimated thickness of the shear layer and the radius of the pin. The expression

\[ f_T^+ = C_1 \left(1 + \frac{\delta}{a}\right)^{C_3} \]

Eq. 15

considers only hypothesis iii, and results in the interpolating trend shown in Figure 5. This expression is simple, accurate, and it becomes a power law both in both asymptotic regimes when the shear layer is very thin or very thick. The recommended values for the constants is \(C_1=0.764\), \(C_2=0.259\) and \(C_3=-0.857\). Constants \(C_1\), \(C_2\) and \(C_3\) synthesize many published results for experimental and numerical simulations in aluminum alloys, carbon steel, stainless steel and titanium.

Figure 7 compares the predictions of scaling corrected with Eq. 15. Most of the systematic error has been eliminated, and only scatter and systematic error from materials properties remain. This figure includes a band of +/-1 standard deviation around the identity line representing a relative error of 12%.

7. Conclusions

The coupled behavior of heat and metal flow around the pin in FSW has been modeled using scaling laws. The model generated closed-form expressions for the shear stress in the metal, the thickness of the shear layer, the maximum temperature, and the volumetric heat generation. The expressions use only known process parameters without the need to measure torque or temperatures for model input.

The comparison with published data indicates that the scaling law for maximum temperature captures the proper trends and orders of magnitude of maximum
temperature, torque and shear layer thickness. Before improvement and considering the range of basic hypotheses, the maximum temperature was overpredicted by approximately 25%; this agreement is remarkable considering that this model was generated from fundamental transport equations and did not involve any adjustment factor.

Among the simplifications imposed on the model, Hypothesis i was always satisfied in all the cases analyzed, and the Peclet number did not significantly influence the estimations of maximum temperature. Hypothesis iii (thin shear layer) was not always satisfied, resulting in a systematic deviation from the asymptotic regime. Accounting for this systematic deviation using scaling concepts resulted in predictions within 12% of measurements for experiments performed by a variety of groups using different equipment and techniques, as well as a broad diversity of alloy systems (aluminum, titanium, arbon steel, and stainless steel).

Looking into the future, it is reasonable to envision FSW references containing formulae for relevant questions such as recommendations for velocities and pin geometry based on the concepts presented here, perhaps the creation of, universal maps of defect generation in FSW. It is hoped that banking on the line of work summarized here, future FSW engineers will be able to consider the feasibility of radically new FSW applications before constructing expensive prototypes.

Figure 7: Comparison between maximum temperature in base material reported in literature and its calibrated estimate[16]
8. References